Imprimé le 25 janvier 2006.

Le Sphinx Développement

Parc Altaïs

74650 CHAVANOD

Téléphone : 04 50 69 82 98

Fax : 04 50 69 82 78

e-mail : contact@lesphinx-developpement.fr

Web : http://lesphinx-developpement.fr

Avant propos

Ce manuel accompagne la nouvelle version du Sphinx que vous venez d'acquérir : Le Sphinx – Version 5. Nous vous en félicitons.

Ce logiciel s'appuie très fortement sur les interfaces et les protocoles de Windows. Vous retrouverez ainsi les habitudes que vous avez déjà acquises par l'expérience de ce système (gestion des fichiers, menus, dialogues, listes déroulantes, raccourcis…). Nous nous sommes efforcés de toujours respecter ces conventions bien établies. C'est pour cela que vous parviendrez très facilement à l'utiliser.

Vous serez également guidé par votre connaissance du travail des enquêtes, des études et de la statistique. Nous utilisons le langage de ces métiers et avons structuré le logiciel par rapport aux grandes étapes d'une étude : questionnaire, saisie, dépouillement, analyse… Notre souci constant est de faciliter votre compréhension devant l'écran. A cette fin, de nombreux commentaires sont affichés pour vous aider à vous situer et à comprendre les menus, boutons de commande, options… Pour les opérations les plus complexes, vous êtes pris en charge par un assistant. Enfin, lorsque le sens d'un article ou d'un bouton vous échappe, il vous suffira d'essayer pour comprendre très vite son utilité.

Pour toutes ces raisons, vous n'aurez aucune difficulté à vous servir de votre logiciel et peut-être pourrez-vous même vous passer du manuel. Mais vous risquez alors de le sous-utiliser ou de vous compliquer inutilement la vie. En effet, toutes les possibilités qu'il offre ne sont pas également visibles ni toujours très compréhensibles au premier abord et vous risquez de passer à côté de fonctions très utiles.

Le premier objectif de ce manuel est de vous faire découvrir tout ce que vous pourrez faire avec le logiciel. A cette fin, une prise en main du logiciel vous donne rapidement un aperçu des possibilités de Sphinx à chacun des stades de votre étude : le questionnaire, la collecte des données, le traitement et la communication des résultats. La partie suivante du manuel vous explique le logiciel en décrivant ses fonctions, ses écrans, ses menus, ses dialogues et les états (tableaux, graphiques) qu'il permet de produire.

Si vous êtes déjà utilisateur du Sphinx, vous pourrez rapidement identifier les nombreuses nouveautés et vous retrouverez facilement vos habitudes. Si vous découvrez ce logiciel, vous comprendrez facilement son organisation et son système de fonctionnement.

Enfin, en vous reportant en dernière partie du manuel, vous trouverez des rappels méthodologiques qui pourront vous être utiles dans la réalisation de vos enquêtes. Ils vous aideront à utiliser toutes les possibilités de votre logiciel pour bénéficier des dernières avancées dans les méthodes de collecte et d'analyse de données.

En vous souhaitant un bon travail.

Le Sphinx Développement

Plan du manuel

Table des matières

Dans la suite du manuel, lorsque cela sera nécessaire, nous mentionnerons par (*) et (**) les fonctions respectivement spécifiques à Plus² et à Lexica.

Ce qu’il faut savoir pour commencer à travailler

1. Installation

Vous venez d'acquérir un des produits de la gamme Sphinx et nous vous en remercions. La procédure d'installation est identique quel que soit le produit que vous avez acquis.

Vous avez un CD-ROM

Insérez le CD-ROM dans le lecteur et attendez que l'exécution automatique se déclenche (de 1 à 10 secondes selon les ordinateurs). Si le programme d'installation ne se déclenche pas (petite anomalie du système sur votre ordinateur), ouvrez l'explorateur de fichiers, affichez le contenu du lecteur de CD-ROM et faites un double-clic sur l'icône **Install**.

Vous avez téléchargé le logiciel

Le fichier que vous avez téléchargé se trouve dans le répertoire de stockage des fichiers téléchargés qui est défini dans le navigateur Internet. Faites un double-clic sur ce fichier.

Licence d’utilisation

Avant de procéder à l’installation de votre logiciel, vérifiez que vous êtes en mode administrateur sur votre poste.

Avec une disquette d'installation

Insérez la disquette dans le lecteur et cliquez sur **Continuer**.

Avec une clef d'installation :

Choisissez le second bouton radio, cliquez sur **Continuer**. Entrez exactement le **Nom d'utilisateur** et la **Clef d'installation** qui vous ont été communiqués par mail ou par courrier. Cliquez sur **Continuer** pour lancer l'installation des programmes sur votre ordinateur.

Installation des programmes

Le programme d'installation détecte la configuration qui vous a été remise :

- Nom du produit
- Options complémentaires
- Nom d'utilisateur
- Nombre et type de la licence
- Date d'expiration de la licence

Vérifier que cette configuration correspond à votre commande et cliquer sur ***Installer.*** Vous pouvez ensuite choisir le répertoire d'installation des fichiers de l'application. Par défaut, le répertoire proposé est ***C:\SphinxV5.*** Nous vous conseillons de le conserver si vous n’avez pas de contrainte d’installation particulière.

Cliquer sur ***Installer.*** Un dernier dialogue vous confirme que le processus d'installation s'est déroulé correctement. L'application est désormais prête. Pour démarrer le programme : un raccourci a été créé sur le **Bureau** ou un menu **Le Sphinx** a été créé dans le menu ***Programmes.***

Validation de la clef

Après installation de votre logiciel, une activation de licence est demandée. Si elle n’est pas réalisée dans un délai de 2 semaines après l’installation du Sphinx, il ne sera plus fonctionnel. L’activation peut être faite par Internet ou par téléphone.

Activation par Internet :

Validez simplement afin d’activer votre licence. Dans le cas où l’activation ne se déroule pas correctement (problème de connexion ou autre), référez-vous à l’activation par téléphone.

Activation par téléphone :

Veuillez nous transmettre vos 2 codes (C1 et C2) par téléphone au 04.50.69.82.98. ou par email à l’adresse contact@lesphinx-developpement.fr.

2. Procédures complémentaires

Mise à jour automatique

Il est possible de vérifier automatiquement si une mise à jour est disponible et de la télécharger pour l'installer automatiquement.

Déclenchez la vérification par le bouton **Mise à jour** présent dans l'écran de démarrage ou par l'article **Mise à jour** du menu **?**.

Déplacer la licence

Avant de commencer la désinstallation de Sphinx, il est recommandé de sauvegarder vos enquêtes. Pour déplacer la licence de Sphinx d'un ordinateur A à un ordinateur B :

* Sur l'ordinateur A :

Cliquez sur **Démarrer/Programmes** et dans **Le Sphinx**, choisissez **Désinstaller ou Réinstaller Le Sphinx-V5**.

Dans le dialogue suivant, choisissez **Désinstaller** et cliquez sur **Continuer**.

Sélectionnez **J'ai une disquette d'installation**, insérez votre disquette dans le lecteur et cliquez sur **Désinstaller**. Si le logiciel n'a plus à être utilisé sur la machine actuellement équipée, cochez la case **Supprimer les programmes**.

* Sur l'ordinateur B :

Insérez le CD-Rom puis la disquette et procédez à une installation complète.

Installer une Licence de site

Si vous avez une licence de site, il est possible de déployer le logiciel sur un ensemble de machines ou d'installer le logiciel sur un serveur d'applications.

Déploiement

Installer le logiciel sur un poste quelconque, et déployer sur chacun des postes le répertoire (C :SphinxV5 par défaut), ainsi que les raccourcis créés.

Installation sur un serveur

L'installation sur le serveur se fait comme une installation standard. On pourra ensuite accéder au logiciel depuis d'autres postes, sur lesquels on crée le raccourci vers le programme.

Ce mode d'installation est à utiliser avec précaution car le Sphinx n'a pas une architecture client-serveur, les temps de réponse pourront être longs. Le mieux est de procéder au même type d'installation que celui des logiciels bureautiques (Word – Excel).

Si chaque utilisateur a une zone de stockage allouée sur un serveur de fichiers, on indiquera cette zone comme répertoire de démarrage de l'icône.

3. Les différents logiciels Sphinx

Les trois niveaux de logiciels

Les logiciels Sphinx existent sous trois formes : Sphinx Primo, Sphinx Plus² et Sphinx Lexica. Chaque logiciel permet de gérer toutes les étapes d'une étude : questionnaire, saisie, traitements quantitatifs et qualitatifs, communication des résultats. Ils se distinguent cependant au niveau :

des traitements par des compétences plus étendues en analyse quantitative pour Plus², qui ajoute notamment des possibilités d'analyse multivariée, et pour Lexica qui approfondit encore les possibilités d'analyse textuelle ;

de l'ouverture de données : Primo est limité aux seules données qu'il permet de saisir, Plus² permet de traiter des bases de données existantes et Lexica ouvre n'importe quel texte.

Le Sphinx Primo

Avec le Sphinx Primo, vous pouvez :

concevoir vos questionnaires, saisir vos réponses et dépouiller vos enquêtes. Du questionnaire au rapport, vous disposez d'un outil pratique et puissant ;

détailler et synthétiser vos données d'enquêtes. Qu'il s'agisse de vérifier les résultats et corriger les données, de préparer et éditer des listes, de rechercher les particularités de certains individus, vous naviguerez très facilement dans vos données. Le Sphinx est un puissant explorateur de données.

Le Sphinx Plus^2

Le Sphinx Plus² intègre toutes les fonctions de Primo et vous permettra en outre de :

approfondir vos analyses grâce aux techniques statistiques avancées de l'analyse multidimensionnelle : analyse factorielle, classification, typologie, analyse de la variance… ;

enrichir votre base de données initiale en calculant de nouvelles variables ;

ouvrir et analyser tout fichier contenant des nombres et/ou du texte. Outil d'analyse statistique ouvert, le Sphinx Plus² vous permet d'analyser toutes vos données quelle que soit leur origine.

Le Sphinx Lexica

Enfin, avec Le Sphinx Lexica, vous pourrez enrichir considérablement les possibilités d'analyse de contenu et d'analyse lexicale déjà présentes dans Primo ou Plus². Vous aborderez ainsi indifféremment les questions ouvertes de vos enquêtes, vos interviews non directives, vos compte-rendus d'entretiens ou tout autre texte quelle qu'en soit l'origine :

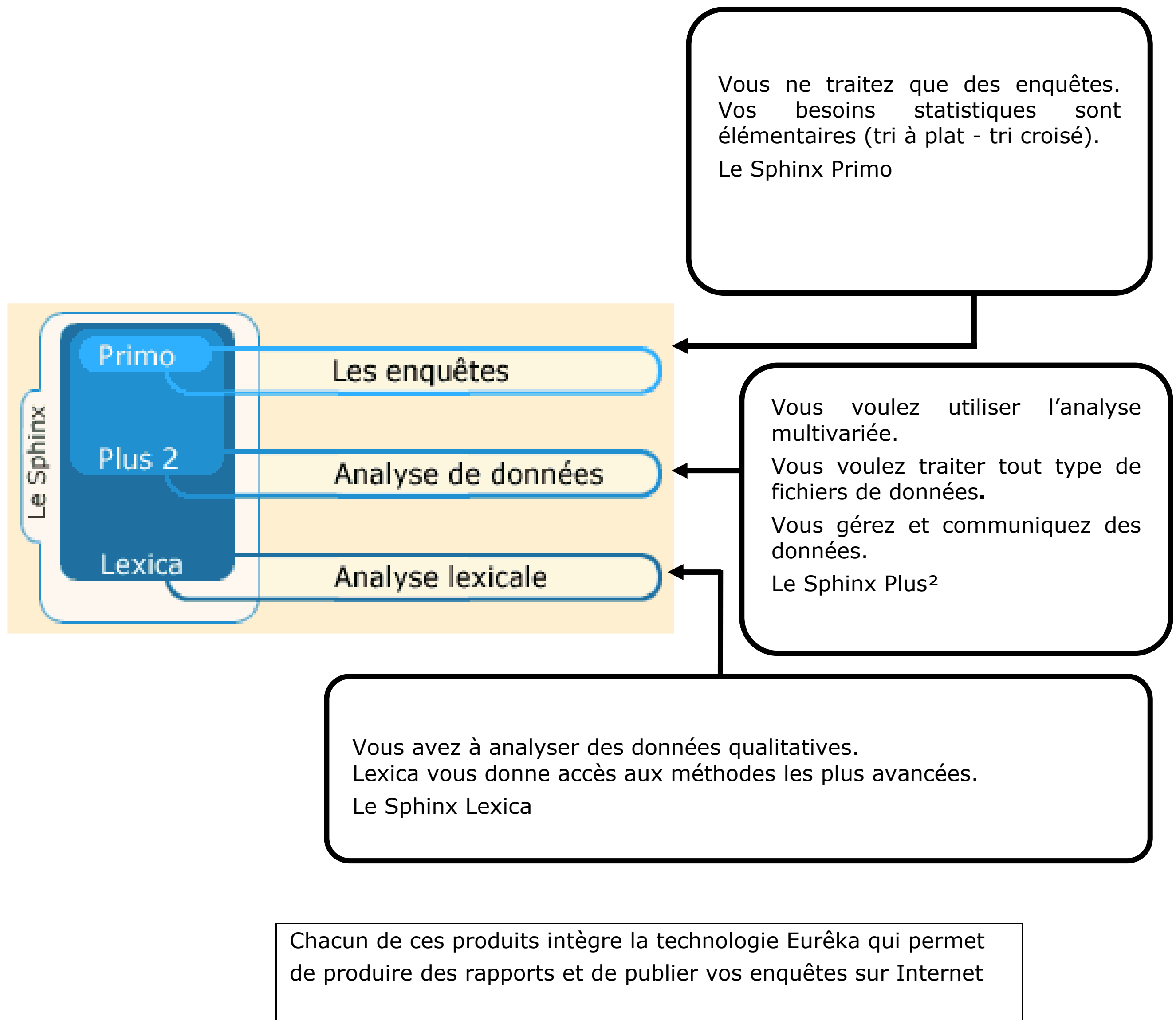
Gérer des listes de mots et d'expressions,

Effectuer des analyses thématiques,

Produire des extraits de textes et restituer des phrases caractéristiques,

Calculer des statistiques lexicales qui permettent de quantifier vos textes…

Dans la suite du manuel, lorsque cela sera nécessaire, nous mentionnerons par (*) et (**) les fonctions respectivement spécifiques à Plus² et à Lexica.



Chacun de ces produits intègre la technologie Eurêka qui permet de produire des rapports et de publier vos enquêtes sur Internet

4. A l'intention des utilisateurs d'une version antérieure du Sphinx

Si vous possédez déjà une ou plusieurs licences des logiciels Sphinx, vous bénéficierez avec la version 5 de :

plus d'expertise dans votre enquête ;

plus d'impact dans la communication de vos résultats ;

plus d'ouverture vers d'autres logiciels ;

une passerelle plus directe vers l'Internet.

Les nouveautés

Les principes et l'esprit restent les mêmes : vous retrouverez les fonctions que vous connaissez déjà.

Avec Windows 95, Windows 98, Windows NT et les nouveaux systèmes, la capacité d'adressage à 32 bits autorise le traitement de données très volumineuses sans sacrifier la rapidité. Le Sphinx Version 5 y gagne une capacité de traitement illimitée et encore plus de vivacité. Les méga-bases s'offrent à vous comme les enquêtes les plus volumineuses comptées en millions d'observations. Mais comme tout se paie, **vous ne pourrez pas les faire fonctionner sous Windows 3.1 !**

Cette nouvelle puissance a aussi permis de donner au logiciel son nouveau look, la matière est plus riche, l'interface plus séduisante mais toujours aussi efficace.

Les nouvelles capacités enrichissent les trois stades : la définition des questionnaires, la saisie des réponses, le traitement des données ainsi que la présentation des résultats.

Parmi les grandes nouveautés intégrées dans la version 5 :

Amélioration des représentations graphiques (cartes géographiques, histogrammes illustrés, affichage des résultats essentiels des résultats) ;

Simplification des procédures d'importation de données externes ;

Diffusion simplifiée des enquêtes sur le web, avec mise à disposition gratuite d'un hébergement de vos enquêtes en ligne sur SphinxOnline (5 000 unités offertes sur SphinxOnline) ;

Fonction de check permettant d'évaluer la fiabilité du questionnaire, des données et des analyses, et correction orthographique automatique.

Conversion et compatibilité vers les anciennes versions de Sphinx

La version 5 de Sphinx Primo, Plus² et Lexica reconnaîtra bien sûr toutes les enquêtes que vous avez pu créer avec d'autres versions de Sphinx (Versions 2 – 2000 - Millenium pour Windows, Version Macintosh). Et à l'inverse, si vous voulez reprendre, dans des versions antérieures de Sphinx, des enquêtes créées dans Sphinx Version 5, utilisez l'article **Enregistrer sous / Convertir l'enquête pour** du menu **Fichier.**

5. Les trois stades de l'enquête

Le Sphinx est organisé en une structure tripartite qui correspond aux trois stades logiques de réalisation de toute enquête.

Elaboration du questionnaire

Utilisez les fonctionnalités de la partie **Questionnaire** pour :

- Accéder à la rédaction des questions ;

- Organiser votre questionnaire ;

- Accéder directement à l'impression rapide et à la publication sur Internet ;

- Evaluer votre questionnaire : check list et conseils ;

- Trouver un lien vers la table des données ;

- Accéder aux formulaires multimédia, pour une mise en page personnalisée

Collecte des réponses

La partie **Collecte des réponses** vous permet de collecter tous les résultats de vos enquêtes avec la saisie de vos réponses, mais aussi leur gestion, leur consultation ou leur modification :

- Saisie rapide des données ;

- Accès aux fonctions de gestion de données (suppression, rassemblement, import…) ;

- Evaluation du jeu de données : check list et conseils ;

- Lien vers la table des données

- Accès au changement de strate.

Traitements et analyses

Qu'il s'agisse de traitements classiques, (tableaux à plat ou tableaux croisés), ou de fonctionnalités plus élaborées, le Sphinx vous permet au stade des **Traitements et analyses** de faire parler vos données et de les rendre intelligibles.

Vous pourrez ainsi non seulement analyser vos résultats, mais également les communiquer de manière efficace :

- Dépouillement automatique des résultats ;

- Possibilité de changement de strate ;

- Conseils proposés pour l'analyse à partir des caractéristiques du questionnaire et des données ;

- Lien vers la table des données ;

- Analyses de données : tris à plat – tris croisés – analyses multivariées.

- Préparation de tableaux de bord ;

- Fonctions d'analyse de données textuelles (pour Lexica).

6. Pour commencer votre travail

Dès votre arrivée dans le Sphinx, vous vous trouvez dans l'une des situations suivantes :

- création d'une nouvelle enquête ;
- reprise d'une enquête déjà créée ;
- recherche d'une enquête ;
- ouverture d'un fichier de données externes.

Créer une nouvelle enquête

Aucun travail n'a été effectué préalablement. Vous allez donc créer un questionnaire de toutes pièces, avant d'en effectuer la saisie et les traitements. Dans cet objectif, cliquez sur **Nouvelle Enquête** de l'écran de démarrage ou sur **Nouvelle enquête** du menu **Fichier** si une autre étude est ouverte.

Reprendre une enquête

Comme dans la plupart des logiciels Windows, le Sphinx propose un raccourci pour reprendre les 4 derniers fichiers ouverts. Dans l'écran de démarrage, il suffit de sélectionner la ligne correspondant à l'enquête souhaitée et de cliquer sur **Reprendre** pour la relancer. Si le Sphinx est ouvert, sélectionnez la ligne correspondant à l'enquête souhaitée dans le menu **Fichier**.

Si votre enquête ne figure pas parmi les 4 propositions, vous pourrez y accéder grâce au bouton **Ouvrir** disponible dans l'écran de démarrage ou par la fonction **Ouvrir une enquête** du menu **Fichier**. Ces options, selon la démarche habituelle de tout logiciel Windows, vous demandent d'indiquer la localisation de votre enquête sur votre poste.

Rechercher une enquête

Si vous avez oublié la localisation précise de votre enquête, aidez-vous des options de recherche disponibles dans Sphinx : l'article **Chercher une enquête** du menu **Fichier** (disponible dès que vous avez ouvert n'importe quelle enquête) vous permet de :

- indiquer le répertoire dans lequel vous recherchez votre enquête (bouton **Parcourir**) ;
- faire apparaître toutes les enquêtes enregistrées dans un répertoire de travail (bouton **Chercher**) ;
- faire apparaître toutes les enquêtes du poste de travail (bouton **Chercher partout**).

Ouvrir un fichier de données externes (*)

Si les données que vous souhaitez analyser dans Sphinx existent déjà dans un autre format pour avoir été créées dans un autre logiciel, vous pouvez bien sûr les récupérer sans avoir à les saisir de nouveau. Vous utiliserez pour cela le bouton **Données externes** de l'écran de démarrage ou la fonction **Ouvrir un fichier de données** du menu **Fichier** (disponible uniquement dans la version Plus²).

Vous accéderez ainsi à l'ouverture d'un fichier de données préalablement enregistré au format Texte (option **Des données formatées**). Vous pourrez également ouvrir un texte à analyser, une base d'images, un tableau d'effectifs, une étude statistique provenant du logiciel SPSS, ou de tout autre logiciel de gestion de données fonctionnant sous un pilote ODBC : Excel, Access, Paradox, D-Base par exemple… On peut également ouvrir des données sur Internet ou télécharger les résultats d'une enquête publiée.

Prise en main du logiciel

7. Elaboration du questionnaire

Créer une nouvelle enquête

- Depuis l'écran de démarrage du Sphinx, cliquer sur **Nouvelle Enquête**. Sélectionner **Créer un nouveau questionnaire** et cliquer sur **Suivant**.
- L'option **Adapter un questionnaire** permet de rédiger son enquête à partir d'un questionnaire pré-existant ou des modèles de questionnaires proposés par le Sphinx.
- Quand vous avez créé un nouveau questionnaire, les fenêtres qui apparaissent ensuite vous permettent de saisir l'**en-tête** : le titre de l'enquête, le nom de l'organisme, la période et d'éventuels commentaires à destination des enquêteurs ou des répondants.
-

Pour vous exercer :

Entrez "Enquête de démonstration" comme titre, puis "Merci de bien vouloir consacrer quelques minutes pour répondre à ce questionnaire" en commentaire. "Institut ABC" est l'organisme et "Janvier 2006" est la période. Cliquez sur Suivant.

En-tête du questionnaire

Titre : Enquête de démonstration

Commentaire : Merci de bien vouloir consacrer quelques minutes pour répondre à ce questionnaire

Organisme : Institut ABC   Période : Janvier 2006

Langue : FR - Français

Afficher l'en-tête   OK   Annuler

L'étape suivante permet de commencer la rédaction, soit en créant successivement toutes les questions, soit (plus rarement) en définissant les grandes parties du questionnaire.

Cliquer sur **Rédiger la première question** puis sur **Terminer**.

Définition des questions

Le menu de définition d'une question est composé principalement :

- du libellé : il s'agit de l'intitulé complet de la question. Ce texte doit être clair et compréhensible. La qualité des informations recueillies en dépend,
- du nom de la variable : c'est le nom abrégé de la question. Il doit être court et explicite. Il permettra de repérer la question dans les listes et dans les autres stades du logiciel,
- du type de la question : celui-ci doit être sélectionné parmi les sept catégories disponibles,
- d'un espace dédié aux modalités de réponses possibles pour les questions fermées.

On passe à la question suivante en cliquant sur le bouton **Nouvelle question.**

Question n° 1

Libellé : Quel est votre sexe ?

Variable : SEXE   Contrôles...   Bibliothèque...

Type : Fermée (unique) / Fermée (multiple) / Fermée (échelle) / Numérique / Texte / Code / Date / heure

Modalités :   Liste...   Barème...

Homme ; Femme

2 modalités

Indiquer les modalités de réponse en les séparant par des points-virgules (;).

Nouveau titre...   Nouvelle question ...   OK   Annuler

Pour vous exercer :

Entrez "*Quel est votre sexe?*" comme libellé, puis "*SEXE*" comme variable. Cochez Fermée (unique) dans le type. Saisissez *Homme* puis *Femme* dans l'espace réservé aux modalités, en les séparant par un point virgule.

Cliquez sur Nouvelle question.

Dans la zone de la 2è question, entrer "*Durant combien de jours êtes-vous partis en vacances l'an dernier ?*" en libellé, "*NB JOURS VACANCES*" en variable. Cochez le type Numérique. Cliquez ensuite sur le bouton Contrôles afin de définir les limites minimum et maximum pour cette variable : 0 en minimum et 100 en maximum par exemple. Cliquer sur OK pour visualiser la liste des questions.

Les questions sont introduites dans la liste les unes après les autres (Menu **Rédiger**, **Nouvelle question** ou double clic dans l'espace vide en bas de liste).

Les options de la fenêtre de définition des questions

Contrôles

Ce bouton permet d'ajouter des consignes particulières ou des conditions pour la saisie des réponses :

Réponses obligatoires,

Limites des valeurs possibles pour les variables numériques,

Question ouverte "autre précisez",

Raccourcis de saisie,

Conditions de présentation,

Variable calculée automatiquement.

Barème

Le bouton **Barème** ouvre une fenêtre permettant de donner à chaque modalité une valeur numérique (ou un poids) qui sera ensuite prise en compte au moment du dépouillement.

Cette fonction s'utilise typiquement dans le cas d'un QCM d'évaluation ou d'un "scoring".

Bibliothèque

Avec le bouton **Bibliothèque,** on accède à des questions qui peuvent être choisies et introduites dans le questionnaire en cours d'élaboration avec le bouton **Remplir.**

Inversement, le bouton **Enrichir** permet d'ajouter à la bibliothèque une question de votre questionnaire en cours d'élaboration.

Liste (des modalités)

Le bouton **Liste** permet de changer l'ordre des modalités. Après une première saisie, il devient **Modifier** et permet d'ajouter de nouvelles modalités.

Le bouton **Echelles** propose des exemples d'échelles qu'il suffit de sélectionner en cliquant pour les introduire dans l'espace dédié aux modalités de réponses.

Déplacement - Duplication - Suppression

Vous pouvez sélectionner l'élément (ou les éléments) à déplacer, en cliquant dans la petite zone à l'extrême gauche de l'écran. Une petite main apparaît et vous pouvez alors faire glisser l'élément vers l'emplacement désiré.

Pour dupliquer ou supprimer des questions ou des titres de partie, la séquence est la suivante : Menu **Rédiger**, **Dupliquer les questions** ou Menu **Organiser**, **Supprimer**.

Pour supprimer une ou plusieurs questions, on peut également la (ou les) sélectionner, puis taper simultanément au clavier sur Ctrl et Suppr.

Question tableau ou groupe de questions

Un groupe de questions est un ensemble de questions successives dont les modalités de réponses sont identiques. Dans les enquêtes de satisfaction par exemple, les répondants se prononcent sur les différentes dimensions du service. Usuellement, cet ensemble de questions leur est présenté sous la forme d'un tableau, appelé dans le Sphinx "groupe de questions".

Dans le logiciel, il existe 2 manières de procéder :

- Créer autant de questions qu'il y a de critères à évaluer, puis les regrouper.
- Créer directement le groupe de questions.

Grouper des questions de modalités identiques

La séquence est la suivante : il faut tout d'abord créer chacune des questions, avec le même type et les mêmes modalités. Les sélectionner ensuite dans la liste des questions, les grouper avec le menu **Rédiger** (ou par un clic droit) puis **Nouveau groupe de questions**.

Créer directement un groupe

Il est plus rapide de créer directement le groupe avec la séquence : **Rédiger**, **Nouveau groupe de questions**. Un menu spécifique est alors disponible pour saisir le titre du groupe (libellé commun à toutes les questions du groupe), son nom abrégé, les différentes questions du groupe (pour lesquelles on indique le libellé et le nom de variable) et enfin, les modalités communes de réponses.

Pour vous exercer :

Au stade de l'élaboration du questionnaire, choisissez Rédiger, puis Nouveau groupe de questions. Dans la fenêtre qui s'ouvre rentrer :

- dans la zone Titre ; le libellé d'introduction de l'ensemble des questions : "*Merci d'indiquer votre satisfaction sur les points suivants*"

- dans la zone Libellé ; le texte présentant chaque question (en remplacement de question x automatiquement affiché). Cliquez sur Ajouter pour créer une nouvelle question (ligne du tableau). Entrez successivement : *le rapport qualité-prix*, *le dépaysement*, *le logement*, *les activités* puis enfin *le contact avec l'habitant*.

- dans la zone Variable ; on saisit ici le nom des variables (par défaut c'est le libellé de la question) : *qualité-prix*, *dépaysement*, *logement*, *activités* et *contact* dans notre exemple.

- rédigez enfin les modalités de réponses dans la zone Modalités. Saisir *Très satisfait ; Satisfait ; Indifférent ; Insatisfait ; Très insatisfait* séparés toujours par des points-virgules. Le bouton Echelles permet de choisir éventuellement parmi la liste d'échelons pré-définis.

La liste des questions est modifiable grâce aux boutons **Monter**, **Descendre**, **Supprimer**, **Ajouter**.

Modification des groupes

Le menu **Rédiger**, **Modifier les groupes** ouvre à nouveau la fenêtre de définition des groupes. En bas à gauche les flèches permettent de parcourir les groupes déjà créés et de les modifier en ajoutant, supprimant ou déplaçant une ou plusieurs questions.

Les boutons **Défaire** et **Tout Défaire** suppriment le groupe sélectionné ou tous les groupes. Les questions restent alors dans la liste mais ne sont plus reliées comme faisant partie d'un groupe.

Présentation des groupes de questions

Au moment de l'édition du questionnaire, à la saisie, puis au dépouillement, les questions reliées dans un groupe seront automatiquement présentées sous la forme d'un tableau, avec plusieurs options de présentation qui seront abordées ultérieurement.

Vérification de l'orthographe

(Non disponible dans la version Primo)

Depuis le menu **Rédiger**, **Vérifier l'orthographe**, on peut lancer le vérificateur orthographique de Word® et l'appliquer aux libellés et aux modalités des questions que vous avez rédigées. Une fenêtre apparaît alors qui vous invite à effectuer ou non les corrections proposées. Cette fonction permet en outre de corriger les fautes dans les réponses textuelles.

Structuration du questionnaire

L'en-tête

L'en-tête contient les informations qui présentent l'étude : **Rédiger** / **En-tête** ou accès direct par le bouton **En-tête**, en haut au centre de la liste des questions. L'en-tête a pu être préalablement défini lors de la toute première phase création de l'enquête.

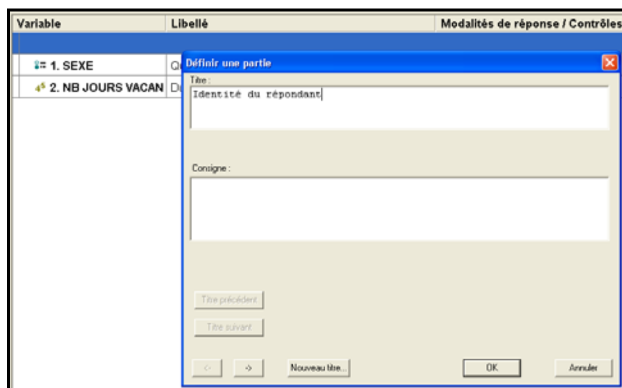
Insertion d'un titre de partie

Les titres de partie permettent de structurer le questionnaire (**Organiser**, **Insérer un titre de partie**). Ils sont insérés avant la question sélectionnée dans la liste et apparaîtront en bleu. Si aucune question n'est sélectionnée, ils sont édités en fin de questionnaire et il faudra alors les déplacer.

Leur contenu est défini dans la fenêtre de définition des titres. Cette fenêtre s'ouvre par un double clic sur l'élément sélectionné ou par le menu **Organiser**, **Nouveau titre de partie**.

Pour vous exercer :

Depuis la liste des questions. Positionner le curseur sur la première question qui est alors surlignée en bleu. Dans le menu Organiser, choisir Insérer un titre de partie. Dans l'espace alors disponible, inscrivez "Identité du répondant" en guise de titre. Cliquez sur OK.



Ces parties sont accessibles au niveau du menu déroulant **Parties**, située en haut d'écran, pour naviguer plus aisément dans un questionnaire très long.

Les renvois (ou sauts)

Les renvois sont utilisés pour sauter des questions et aller plus loin dans le questionnaire en fonction de la réponse à une question filtre. Pour les programmer, on utilise la fonction **Organiser** / **Définir des renvois** et l'on procède comme suit :

1/ On choisit d'abord la question filtre depuis : **Après la question...**

2/ On fixe les conditions du renvoi en cliquant sur : **Définir le profil...** en indiquant les modalités ou la plage de valeurs qui déclenchent le renvoi. Un renvoi peut être simple ou multiple. Dans ce dernier cas, on introduira plusieurs conditions en cliquant sur le bouton **Ajouter.**

3/ On indique la cible avec : **Aller à la question...** ou **Terminer le questionnaire**

Pour vous exercer :

Ouvrir l'enquête Démo dans C:\SphinxV5\Enquêtes

Dans l'élaboration du questionnaire choisir le menu Organiser / Définir des renvois.

Dans le menu déroulant Après la question, choisir la question n°2 (CSP).

Cliquez sur Définir le profil. Dans la liste des modalités, choisir Inactifs, autre puis cliquez sur OK.

Dans le menu déroulant Aller à la question, choisir la question n°8 (Activités préférées). Cliquez sur Fermer.

On revient alors sur la liste des questions. Les crochets rouges et roses rappellent les points de départ et d'arrivée des renvois.

Lorsqu'ils sont activés, les renvois se déclenchent automatiquement à la saisie.

Attention : sur Internet, ils ne peuvent fonctionner que si la question filtre et la cible sont sur des écrans différents.

Avec le bouton **Ajouter**, on peut définir autant de renvois que nécessaire. Ils s'ajoutent alors dans la liste des renvois.

Attention : Il est déconseillé de créer trop de renvois : le répondant peut s'y perdre et vous aussi !

Les conditions de présentation

Dans les formulaires multimédia, ces conditions permettent de faire apparaître une question selon qu'une (ou plusieurs) condition(s) soi(en)t remplies. On y accède par le bouton **Contrôles** lorsqu'on est sur l'écran de la question à afficher sous condition, ou par **Organiser** / **Définir des conditions**…

Renvois et conditions de présentations peuvent être combinés, offrant ainsi de nombreuses possibilités d'interaction dans votre formulaire.

Evaluation du questionnaire

Au stade **Elaboration du questionnaire**, vous avez accès à l'évaluation du questionnaire. Le Sphinx vous demande tout d'abord de spécifier le contexte de votre travail afin de vous renvoyer les conseils les mieux adaptés aux caractéristiques de votre enquête :

- le **contexte de l'étude** : son objectif, le mode de collecte, la cible et le nombre de réponses attendues,
- le **contexte de la collecte** : le nombre de questionnaires diffusés (ou de contacts pris), la nature de l'enquête (sondage ou enquête exhaustive), les éventuels quotas à respecter,
- le **contexte de l'analyse** : les questions-clé qui seront particulièrement importantes dans les analyses à effectuer.

En fonction des éléments que vous aurez déclarés, le logiciel vous propose un diagnostic portant sur :

- le nombre de questions, leurs types et leur formulation,
- la complexité du questionnaire,
- le temps estimé de saisie.

Production automatique du questionnaire, papier ou écran

A partir de la liste des questions, on obtient un questionnaire prêt à l'emploi soit en l'imprimant dans l'environnement classique soit en le mettant sur Internet depuis l'environnement multimédia.

Dans l'environnement classique

Panneau de commande / **Elaboration du questionnaire** / **Impression rapide**.

Cette séquence lance la mise en forme du questionnaire qui s'affiche dans la fenêtre Impression des questionnaires.

Mise en forme et impression

Le bouton **Mise en page** permet de modifier les règles de mise en forme (disposition sur une ou plusieurs colonnes, contenu de l'en-tête, numérotation et présentation des questions, choix des polices, type d'affichage des réponses). Les options paramétrables dans les différents onglets offrent de nombreuses possibilités et permettent d'obtenir un document opérationnel très rapidement.

La coche **Conserver les paramètres** permet d'enregistrer vos modifications et de les retrouver à la prochaine mise en page automatique.

Le bouton **Imprimer** lance l'impression.

Attention : l'orientation du papier se définit dans le dialogue avec l'imprimante (**Configuration de l'imprimante**).

Dans le cas d'une impression automatique par le menu **Fichier**, **Imprimer,** l'option de mise en page "**Modalités en tête de colonne pour les groupes de questions fermées**" permet de libeller en clair les titres de colonnes. Si la case n'est pas cochée, un numéro renvoie à une légende située au bas du tableau. Cette option est accessible par **Fichier**, **Mise en page** et l'onglet **Numérotation**.

Transfert dans un traitement de textes

Fichier / **Inclure dans le rapport** lance Word® dans lequel votre questionnaire est copié conformément aux options de mise en forme choisies dans **Mise en page**.

Mise en forme rapide des formulaires multimédia

Depuis le panneau central, au stade **Elaboration du questionnaire**, on clique sur **Formulaires multimédia**.

Dans l'assistant, choisissez **Papier** puis faites le choix d'un **Style** (une colonne ou classique ou condensé...) puis **Terminer**.

Votre questionnaire apparaît conformément au style choisi. Optez alors pour :

- **Fichier** / **Imprimer** pour l'imprimer,
- ou **Fichier** / **Enregistrer dans Office** pour reproduire le formulaire dans Word® ou PowerPoint®.

Mise en page personnalisée de formulaires papier ou écran

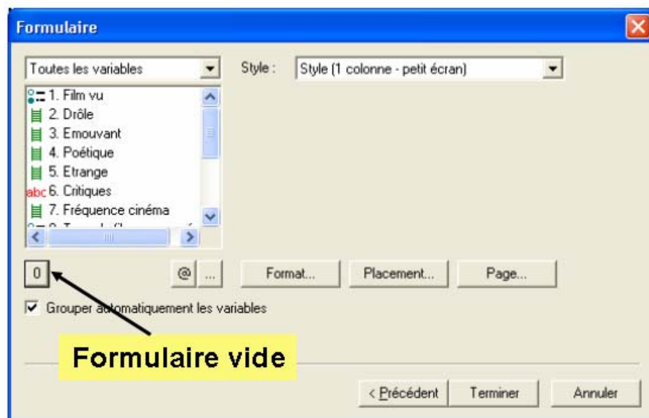
On procède ainsi lorsqu'on veut :

- Contrôler le nombre de pages,
- Appliquer une charte graphique,
- Ajouter des légendes ou des illustrations,
- Faire des formulaires interactifs,
- Préparer un script pour une enquête téléphonique,

…. et toute autre mise en page adaptée à chaque médium et à chaque circonstance particulière.

Depuis le stade **Elaboration du questionnaire** du panneau de commandes central, on opte pour **Formulaires multimédia** / **Formulaire** / **Nouveau formulaire.**

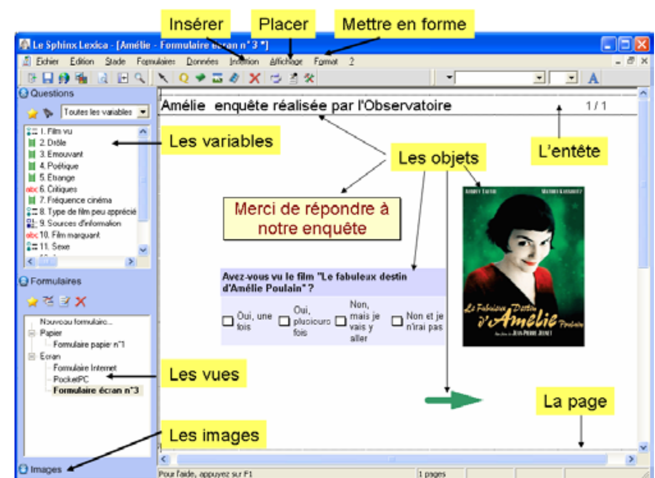
On crée une vue vide en désélectionnant toutes les variables (bouton 0). Les boutons **Format**, **Placement**, **Page**, permettent de choisir les propriétés de la nouvelle vue et des objets qui y seront insérés.



L'environnement de création du formulaire

Cet environnement est constitué d'une vue divisée en pages dans laquelle on compose le formulaire en plaçant les questions, les illustrations, les légendes et éventuellement les boutons d'actions...

A gauche, la fenêtre du haut donne la liste des variables (questions). Au dessous figure la liste des formulaires déjà créés. On voit enfin la liste des images.

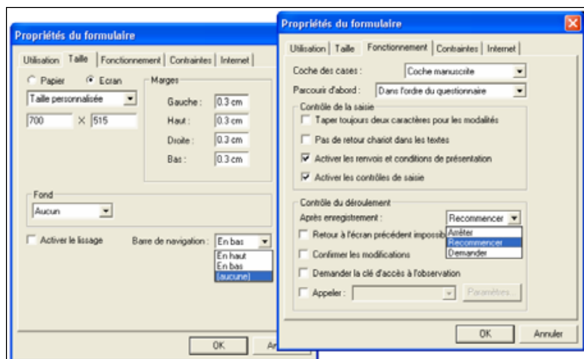


De nombreuses techniques sont disponibles pour placer des objets dans la vue :

- en cliquant dans la liste des variables pour ajouter une question (cliquer glisser ou double clic),
- en cliquant dans la liste des images pour ajouter une illustration,
- à partir du menu **Insertion** pour une question, un groupe de questions ou toutes les questions d'une partie,
- à partir du même menu **Insertion** pour une légende, une image, une forme graphique,
- en utilisant les raccourcis accessibles par la barre d'outils.

La vue et ses propriétés

On appelle "vue" l'ensemble de pages dans lequel on construit le formulaire. On en fixe le format et le mode de fonctionnement dans le panneau **Propriétés du formulaire**. Pour y accéder, vous avez plusieurs possibilités : la fonction **Formulaire** / **Propriétés du formulaire** ou le raccourci situé en tête de la fenêtre des formulaires ou le menu contextuel (un clic droit) dans une zone vierge de la vue.



Dans l'onglet **Taille**, on peut :

- Fixer les dimensions de la page, ses marges et la couleur ou l'illustration de fond,
- Choisir le mode de navigation pour les formulaires Internet (barres de navigation),
- Optimiser l'affichage graphique avec la case **Activer le lissage** qui permet notamment de traiter les superpositions par transparence.

L'onglet **Fonctionnement** permet de régler le mode de saisie, avec notamment l'activation des renvois, des contrôles de saisie ainsi que l'ordre selon lequel les questions doivent être parcourues.

L'onglet **Contraintes** permet d'exprimer les "contraintes d'intégrité" de l'enquête, qui sont en fait des conditions générales que devront respecter les données saisies, pour être cohérentes entre elles. On peut par exemple contrôler le fait qu'une personne déclarant avoir moins de 16 ans, ne puisse pas posséder une automobile...

La construction pas-à-pas : formats et alignements

Pour définir le format des questions en "pas à pas", on procédera de différentes manières possibles:

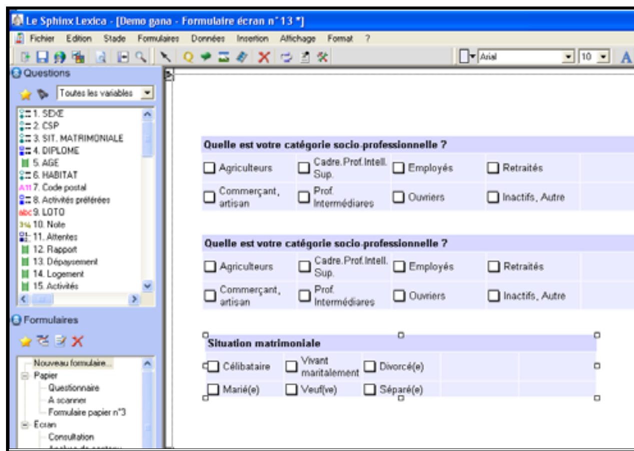
- en transférant le format d'une première question sur une autre ou sur toutes les autres (**Format** / **Copier le Format** puis clic droit vers le menu contextuel **Coller le Format**),

- en définissant le format d'une première question comme le format standard à appliquer à toute nouvelle question insérée dans le formulaire (**Choisir comme Format**),

Il est toujours intéressant de respecter certaines règles plutôt que de chercher des dispositions trop singulières. Ceci permet de procéder partie par partie (**Insertion** / **Partie**). Les règles de placement (**Affichage** / **Placer**) donnent en général un résultat que l'on pourra modifier lorsque cela est nécessaire, tout en prodiguant au formulaire une certaine homogénéité.

Pour un rendu précis, il convient de bien aligner les objets (**Affichage** / **Aligner** / **Répartir** ou les raccourcis Ctrl + flèches). Pour accélérer la construction du questionnaire, on aura ensuite intérêt à fixer l'aspect des différents objets pour les définir comme un "style" (menu **Format / Définir les styles**).

> Pour vous exercer :
>
> Ouvrir l'enquête Démo dans C:\SphinxV5\Enquêtes. Choisir Formulaires multimédia
>
> - A placer des questions
>
> Cliquer sur Nouveau formulaire au centre à gauche. Choisissez Internet puis cliquez sur Suivant. Cliquez sur 0 puis Terminer pour arriver sur un formulaire vierge.
>
> Sélectionner la variable n°2 et faites la glisser (tout en maintenant votre clic) sur la page blanche. Faites de même pour les variables n°3 et 4. Placez-les les unes en dessous des autres sur la première page. Faites glissez les variables n° 5, 6 et 7 sur la deuxième page disponible.



Pour vous exercer :

- A modifier le format des questions

Double cliquez sur la question n°2 pour en modifier les propriétés. Dans l'onglet contenu, cochez Numéro. Dans l'onglet Zones, choisir une forme ronde. Dans l'onglet Affichage, choisir pour le libellé une couleur de fond orange clair. En cliquant sur …, passez la taille de la police de caractère à 12. Cliquez sur OK.

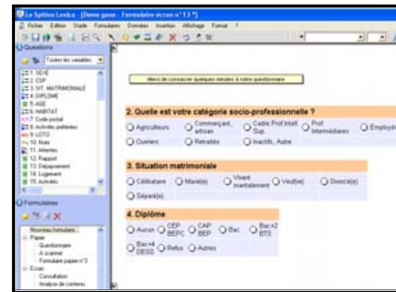
Faire un clic droit sur la question 2 et choisir Copier le format. Taper Ctrl + A pour sélectionner toutes les questions de la vue puis clic droit pour coller le format. Dans le menu Affichage, choisissez Aligner/Répartir puis A gauche. Toutes les questions ont désormais la même présentation et sont alignées à gauche.

- A insérer une image

Choisir le menu Insertion puis Image. Sélectionner le logo du Sphinx dans la catégorie Multimédia – Images. Placez-le en haut à droite de la première page. Faites un clic droit sur le logo et optez pour Choisir comme masque : le logo apparaîtra sur toutes les pages au même endroit.

- A insérer une légende

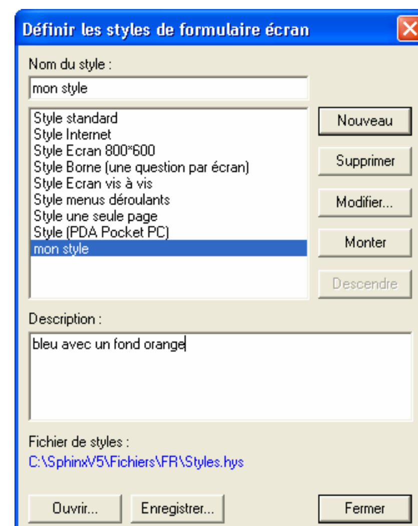
Double-cliquez sur une zone vierge du formulaire et saisissez : "Merci de consacrer quelques minutes à notre questionnaire".



Définition des "styles" de présentation

Le style permet d'enregistrer les formats de présentation souhaités (couleurs, polices, format des coches etc.) pour les appliquer à de nouveaux formulaires qui seront créés ultérieurement. Le Sphinx vous propose des styles de présentation par défaut mais vous pouvez créer vos propres styles.

Pour cela, une fois votre présentation réalisée sur un formulaire donné, vous pouvez enregistrer votre style. Choisir le menu **Format** puis **Définir les styles**. Cliquer alors sur **Nouveau**, donner un nom à son style et éventuellement une **Description**.



8. La collecte et la saisie de réponses

La saisie (collecte des réponses) est le 2ème stade de la réalisation d'une enquête.

La saisie des données au clavier

Ce type de saisie (au clavier) ne concerne que les enquêtes papier.

Plusieurs possibilités sont envisageables :

- la **Saisie rapide** : la plus utilisée, avec les numéros des modalités saisies "au kilomètre"
- la **Saisie directe** : saisie moins rapide, question après question,
- le **Module opérateur** (saisie multimédia) : pour personnaliser la saisie.

Saisie rapide

Dans ce mode de saisie, une observation complète figurera sur une seule page.

Cette fonction permet de saisir vos réponses de manière simple et efficace, avec les numéros des modalités saisies sur un clavier numérique ou cliquées dans la liste des modalités à droite de l'écran.

Pour les questions fermées uniques ou échelles, on passe directement à la ligne suivante après saisie de la réponse (quand il y a moins de 10 modalités).

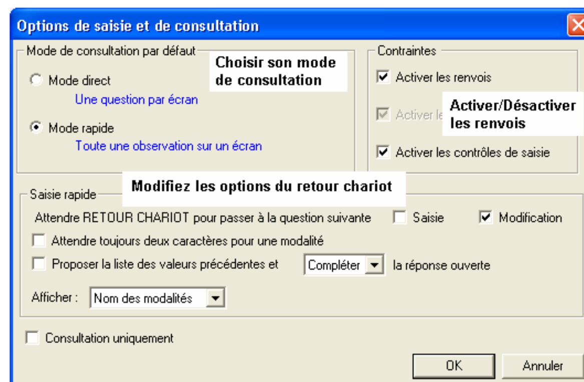
Le bouton **Compléter** met fin à votre saisie par des non-réponses. Ce mode de saisie offre en outre plusieurs options.

Au stade **Collecte des réponses**, menu **Saisir**, les **Options de saisie** permettent de modifier le mode de consultation par défaut.

Vous avez également la possibilité d'activer ou de désactiver les contraintes définies lors de la création du questionnaire (renvois, restrictions et contrôles de saisie).

Le retour chariot quant à lui peut être activé en saisie ainsi qu'en modification.

Attention, son activation influe sur la saisie de questions ouvertes et la gestion des retours à la ligne !



Pour vous exercer :

Depuis le stade Collecte des réponses, cliquez sur Saisie rapide. La liste des variables apparaît à gauche de l'écran, la zone de saisie en haut à droite, au dessus des modalités de réponse.

Tapez 1 au clavier, pour saisir la réponse d'un homme. On est passé directement à la question suivante. Tapez 8 pour Inactifs-Autre à la question n°2 CSP. On est passé à la variable n°8, à cause du renvoi défini précédemment. Pour la question n°8, saisir les modalités 1 puis 2, puis Entrée afin de passer à la question suivante. La variable 9 est une question texte, saisir le texte puis Entrée pour passer à la question suivante. Etc.

Après la saisie de la dernière question, on passe directement à une nouvelle observation.

Saisie opérateur

Le module opérateur permet la saisie de vos données dans un environnement plus agréable mais sans doute moins efficace sur le plan de la rapidité.

La mise en place du module opérateur nécessite, dans un premier temps, la création d'un Formulaire (**Elaboration du questionnaire**, **Formulaires Multimédia**). Parmi les différents types de formulaire, il est recommandé de sélectionner le formulaire écran ou Internet, pour une présentation adaptée à un écran d'ordinateur personnel.

Dès lors que le formulaire est créé et sauvegardé, vous pouvez faire la saisie de vos questionnaires à l'aide du module opérateur (**Collecte des réponses**, **Module Opérateur**, **Saisir**). Si plusieurs formulaires ont été créés, il vous faudra préciser le formulaire sur lequel la saisie est réalisée.

Le module opérateur existe également sous une version réseau permettant à plusieurs postes de travail de réaliser la saisie de vos enquêtes, sans pour autant être équipé du logiciel Sphinx complet ! Enfin, outre les possibilités de saisie, le module opérateur facilite également la communication de vos résultats (sous forme de tableaux de bord).

Programme minimal de saisie ou "visionneur" (*)

(Non disponible dans la version Primo)

Le programme minimal de saisie représente une autre possibilité de saisie délocalisée. Il permet de transmettre une interface de saisie légère aux répondants, qu'ils pourront ouvrir depuis leur ordinateur non équipé de Sphinx.

Son fonctionnement est identique au module opérateur, mais ses possibilités sont toutefois limitées du fait de la faible taille de l'application créée. On constitue un fichier exécutable que vous pouvez envoyer par mail à vos correspondants. Le fichier texte des réponses vous revient en retour.

Cette procédure convient parfaitement à un envoi sur quelques dizaines de répondants. Elle présente toutefois 2 inconvénients. Les fichiers ".exe" peuvent ne pas parvenir au destinataire (protections…).

Par ailleurs, il vous faudra importer tous les fichiers en retour un à un ou bien les compiler dans un seul fichier (par des copier-coller). Depuis les tableaux de bord, l'importation se fait depuis le menu **Données**, puis la rubrique **Importer**, où l'on choisit Importer des réponses depuis **Un fichier de réponses**, dans le menu déroulant.

Le programme minimal de saisie se crée depuis les formulaires ou tableaux de bord selon la séquence : **Fichier**, **Emporter l'enquête**, **Programme minimal de saisie**.

Consulter/Modifier

On y accède depuis le stade **Collecte des réponses** puis **Consulter/modifier**.

Cette fonction est utilisée pour consulter la base de données, pour apporter des modifications aux réponses que vous avez précédemment saisies ou pour effectuer des remplacements.

Consultation et modification peuvent s'effectuer dans trois environnements différents :

- le mode rapide (identique à la saisie rapide)
- le mode direct (question par question)
- le mode tableur que nous détaillerons ci-après.

On passe d'un environnement à l'autre en cliquant sur le bouton en bas à gauche ou par le bouton **Tableur.**

Mode rapide

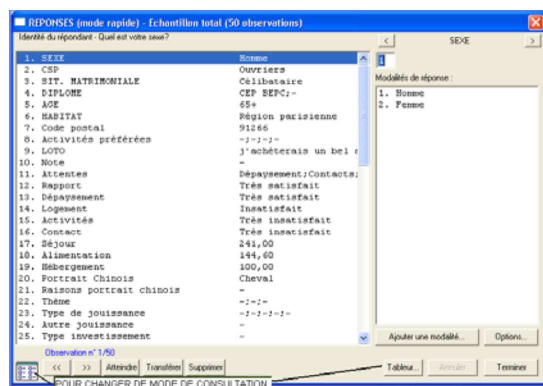
Le mode rapide offre les options suivantes :

- Atteindre : pour se rendre sur une observation précise (suivant son numéro, la valeur d'une modalité, etc.),

- Transférer : pour imprimer les données, créer un rapport, etc,

- Supprimer : pour supprimer l'observation courante,

- Flèches de navigation : pour passer d'une observation à une autre.



Mode tableur

Ce mode permet :

- de visualiser la liste des variables,
- de voir la liste des observations sur une (ou plusieurs) ou toutes les variables,
- de consulter les réponses, fiche par fiche,
- d'éditer des listes de réponses (liste de valeurs).

Depuis la liste des observations, on peut faire apparaître une ou plusieurs variables, grâce au menu déroulant en haut à gauche ou en saisissant le n° des variables dans la zone attenante. Cliquez sur le bouton vert pour obtenir les variables souhaitées. En cliquant sur le nom de la variable en tête de colonne, on peut trier les individus selon les réponses à cette question. En cliquant sur l'icône en haut à droite, on peut imprimer ou copier les données pour les transférer vers un autre logiciel.

Strates | Variables | Observations | Fiches | Liste de valeurs

Toutes les variables | 1.5.7 | Noms des modalités

N°	1. SEXE	5. AGE	7. Code postal
50	Homme		
22	Homme	25-34	99849
5	Homme	18-24	98887
20	Femme	25-34	98662
16	Femme	25-34	97655
3	Femme	65+	94772
13	Homme	65+	94311
46	Femme	35-49	93212
1	Homme	65+	91266
40	Femme	18-24	89789
49	Femme	25-34	88188
23	Homme	35-49	86427
14	Homme	35-49	85864
43	Femme	25-34	82356
6	Homme	25-34	78289
41	Homme	25-34	76336
39	Femme	65+	72982
12	Femme	35-49	72511
11	Homme	18-24	72317
19	Homme	18-24	69252
28	Homme	18-24	66695
35	Homme	50-64	62159
30	Femme	18-24	59473

Remplacer des valeurs

Depuis **Gérer** / **Remplacer des valeurs** ou depuis le bouton **Remplacer** du **Tableur**, on a accès à des fonctions de recherche et de remplacement. Ces opérations peuvent s'effectuer sur une ou toutes les variables, pour toutes les observations ou seulement sur une partie d'entre elles (profil ou plage de numéros). **Attention** : cette action est irréversible !

Rassembler ou fusionner plusieurs fichiers de réponses

Cette série de fonctions est dédiée à la compilation de plusieurs fichiers de réponses au format Sphinx. Il s'agit de fichiers identiques saisis sur plusieurs postes (fonction Rassembler) ou de fichiers différents ayant des variable(s) commune(s) (fonction Fusionner).

Rajouter des observations d'une enquête identique (saisie multi-postes)

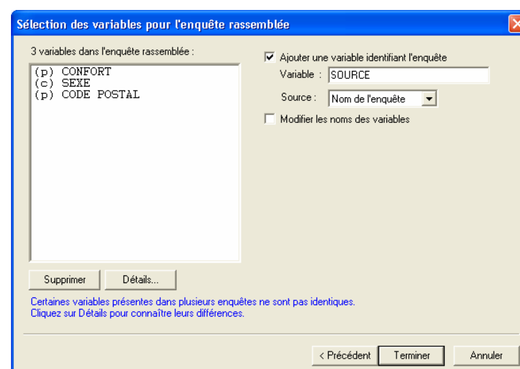
On utilise cette fonction quand des réponses ont été saisies sur plusieurs postes de travail. On indique alors le chemin d'accès de ces fichiers de réponses, dont la terminaison est "nom**.rep**". On ajoute les fichiers correspondant aux données saisies localement. Bien contrôler le nombre total de réponses grâce aux informations disponibles en bas à gauche du panneau de commandes central : nombre de questions et surtout nombre de réponses. **Attention** : Les fichiers de réponses à rassembler doivent être rigoureusement identiques : nombre de questions, noms et types des variables, nombre de modalités.



Rassembler les variables communes de plusieurs enquêtes (*)

Il s'agit de regrouper plusieurs enquêtes ayant des variables communes. On crée alors une nouvelle enquête issue de cette consolidation, qui inclut les questions communes à toutes les enquêtes.

On peut y faire figurer également des variables "partiellement" communes, qui sont présentes dans au moins deux études, mais pas dans toutes. Un menu permet de choisir les variables à fusionner et vous indique si elles sont communes (c) ou partiellement communes (p).



Fusionner deux enquêtes (*)

Par cette fonction, on va pouvoir ajouter à l'enquête en cours des informations concernant les mêmes individus et qui sont contenues dans une autre enquête. Une variable "**clef**", commune aux deux enquêtes, est nécessaire pour identifier le même individu dans les 2 fichiers. Un menu permet d'indiquer la **clef** et les variables à prendre en compte dans la 2è enquête.



Mise en ligne et diffusion des formulaires sur l'Internet

A partir d'un formulaire multimédia correctement mis en forme pour l'Internet (voir chapitre "Mise en page personnalisée de formulaires papier ou écran"), choisissez le menu **Fichier** / **Publier sur Internet** ou activez l'icône Internet situé à gauche de la barre d'outils.

Publication rapide

Pour vous exercer :

Choisir Fichier / Publier sur Internet

Dans l'assistant :

1°/ Cochez : Passer les options avancées

2°/ Indiquez où vous souhaitez publier votre enquête en entrant le nom et mot de passe de votre compte SphinxOnline. Vous pouvez aussi publier sur votre propre serveur d'enquête.

Enfin, pour vos essais, le compte gratuit est à votre disposition (choisir "Accès gratuit").

3°/ L'assistant confirme les choix que vous avez faits et lance la publication. Les fichiers de votre enquête sont alors transférés vers le serveur.

4°/ En retour, vous recevez les liens vers le questionnaire, les tables de données, et éventuellement les résultats si vous avez déjà préparé un tableau de bord.

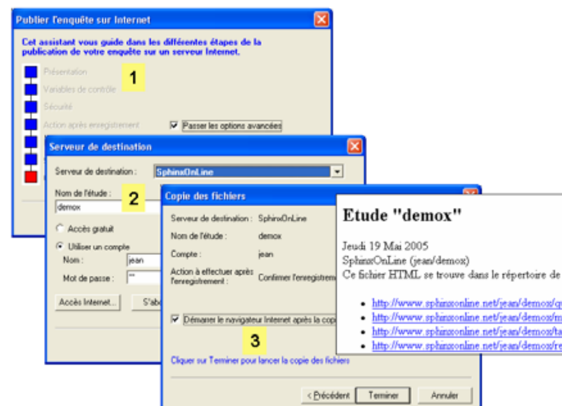
Recopiez le lien vers le questionnaire dans un mail que vous enverrez à vos correspondants. Vous pouvez aussi placer ce lien sur un site.

Si vous avez déjà publié les fichiers de cette enquête, l'assistant vous en avertit et vous propose soit de :

- **Réinitialiser les fichiers sur le serveur** : **Attention** : les réponses déjà recueillies seraient alors perdues !

- **Conserver les réponses du serveur** : ceci n'est possible que si vous n'avez pas modifié la structure de votre questionnaire,

- **Publier les réponses locales** (les données présentes sur votre poste de travail seront transférées sur le serveur). Cette possibilité permet par exemple, en utilisant le contrôle d'accès, de rappeler aux répondants des informations dont on dispose déjà.



Fonctions avancées de publication (*)

Pour avoir accès à ces différentes options, au moment de la publication, il vous faut obligatoirement décocher la case **Passez les options avancées** et vous aurez accès à l'assistant de publication :

Présentation : Java ou HTML ?

- **Java** : la page Internet ressemblera exactement à celle préparée dans votre formulaire multimédia. Il faut disposer d'un navigateur Internet correctement configuré (compatible avec l'applet Java de Sphinx).

- **HTML** : la présentation de certaines pages peut être simplifiée par rapport à celle des vos formulaires multimédia. Par ailleurs, la taille des pages s'en trouve allégée.

Nous vous conseillons de publier vos formulaires en format HTML pour vous assurer que tous les répondants pourront accéder à votre enquête !

Variables de contrôles : clé (*)

Dans le cas de questionnaires très longs et /ou se déroulant sur des périodes distinctes, vous avez la possibilité de faire revenir les répondants sur le questionnaire qu'ils ont commencé de saisir.

Pour cela, il vous faut utiliser une clé (**Publier l'enquête sur Internet**, **Variables de contrôle**, **Clé**). Cet identifiant est communiqué au répondant lorsqu'il stoppe sa saisie et doit être à nouveau entré pour que celui-ci puisse se reconnecter.

Identificateurs de saisie (*)

Il est possible de définir une liste d'**identificateurs de saisie**, que l'on transmettra aux répondants, afin de savoir qui a répondu, qui l'on doit relancer.

Sécurité (*)

Les différents paramètres de cette option vous permettent de rendre libre, de restreindre ou d'interdire la saisie, la consultation, l'accès aux résultats aux répondants de votre enquête.

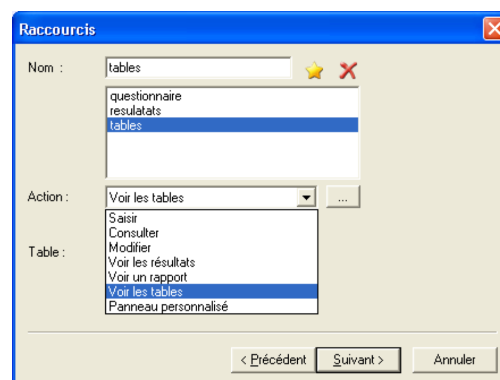
Action après l'enregistrement (*)

Par défaut, le bouton de validation de la saisie se nomme **Enregistrer**. Vous pouvez modifier cet intitulé (ex : valider vos réponses).

Après l'enregistrement des réponses (étape obligatoire), vous pouvez choisir de **confirmer l'enregistrement**, de **rediriger** vers un autre lien Internet (par ex. celui du site de votre entreprise, université, …) ou de **renvoyer des informations** (afficher les résultats par exemple...).

Raccourcis (*)

Au cours de cette étape, vous indiquez les différents liens vers les formulaires, tables et/ou tableaux de bord que vous souhaitez publier. Dans ce stade, vous définissez le **nom** du lien, c'est à dire sa terminaison (par défaut : questionnaire, tables, résultats), l'**action** correspondante dans un menu déroulant, son **format** (java ou HTML), et enfin le formulaire, la table, le tableau de bord ou le rapport à présenter avec le lien correspondant.



Téléchargement des réponses saisies en ligne

Depuis les tableaux de bord, l'importation se fait depuis le menu **Données**, puis la rubrique **Importer**, où l'on choisit **Importer des réponses depuis Sphinxonline** (ou adresse de votre propre serveur d'enquêtes si l'on sélectionne **Autre serveur d'enquêtes**), dans le menu déroulant. Dès lors, le nom de l'étude doit être indiqué ainsi que le nom du compte et son mot de passe.

Attention : Un menu déroulant **Réponses précédentes** permet de préciser si l'on souhaite **Compléter** ou **Remplacer** les réponses précédemment importées ou saisies localement.

Autres modes de saisie

Saisie par scanner (*)

Le scanner permet la lecture des questionnaires de votre enquête. Il s'agit d'un module additionnel au Sphinx (non livré avec le logiciel). Son utilisation intervient au moment de la saisie des réponses et présente plusieurs avantages :

- il évite la saisie manuelle des réponses obtenues (phase qui peut être longue et fastidieuse) et offre un gain de temps important ;

- celui-ci ne mobilise pas de ressources humaines sur une longue période (gain financier et humain)

Si le scanner est utilisé au moment de la saisie, une démarche particulière a dû être adoptée, obligatoirement, au moment de la création du formulaire : celui-ci doit être de type "scanner" (**Elaboration du questionnaire**, **Formulaires multimédia**, **Scanner**).

Attention : une fois le formulaire imprimé sur papier, aucune modification de celui-ci ne doit être apportée dans le logiciel Sphinx. Dans le cas contraire, le scanner ne reconnaîtra plus les formulaires !!! La saisie devra alors se faire manuellement…

La lecture des questionnaires s'effectue automatiquement et les cases cochées par les répondants sont reconnues par le scanner (à l'exception des questions ouvertes).

Attention : pour se servir de la saisie scanner, il est nécessaire d'utiliser un scanner compatible avec le logiciel ! (Pour connaître la liste des scanners compatibles avec Le Sphinx, nous contacter).

Enquêtes CATI, réseau d'opérateurs ou PDA (*)

Le Sphinx permet de déployer la saisie des réponses sur des média très divers :

- un module complémentaire (**Sphinx CATI**) est disponible pour gérer vos enquêtes téléphoniques : partager une base d'appels commune entre plusieurs opérateurs, programmer des rappels, gérer des quotas, suivre en direct les statistiques des appels et les résultats de la campagne.
- depuis les formulaires multimédia, vous créez des pages qui pourront être installées sur des PDA (**organiseurs**), pour des opérateurs de saisie itinérants (salons, transports en commun par ex.), via la fonction **Fichier** / **Emporter l'enquête**/ **Programme de saisie pour Pocket PC**.
- le Sphinx vous permet également de déployer un questionnaire en saisie multipostes (**réseau d'opérateurs**). Les personnes qui effectuent la saisie doivent être équipées d'un module opérateur ("mini-Sphinx" limité à la saisie et à la consultation). Le questionnaire se prépare également en formulaire multimédia. On utilise ensuite la fonction **Fichier** / **Emporter l'enquête** / **Répertoire pour le module opérateur**. Les différents fichiers de réponses pourront être rassemblés ensuite depuis la fonction **Gérer – Rassembler/Fusionner.**

Evaluation des données

Au stade **Collecte des réponses**, vous avez accès à l'évaluation des données qui vous permet d'obtenir une mesure objective de la qualité des réponses que vous avez récoltées.

Toujours en fonction des éléments que vous avez préalablement déclarés, le logiciel vous propose un diagnostic portant notamment sur :

- le calcul des taux de retour et de remplissage,
- le repérage des questions insuffisamment documentées (valeurs manquantes) et des individus présentant des valeurs extrêmes ou aberrantes,
- la caractérisation des données en termes de concentration, dispersion et richesse,
- le décalage entre les réponses obtenues et d'éventuels quotas,
- l'erreur d'estimation associée à votre enquête, s'il s'agit d'un sondage.

Redressement et aide à l'échantillonnage

Depuis le stade **Collecte des réponses**, menu **Gérer**, on a accès à des fonctions de redressement et d'aide à l'échantillonnage (**Redresser** / **Définir des quotas**).

Redressement

Trois méthodes de redressement sont disponibles dans Sphinx, à partir des **quotas** souhaités que vous pouvez spécifier, sur une, deux ou trois variables.

Les méthodes par **substitution** et **extraction** créent une nouvelle enquête redressée. Tous les traitements que vous allez réaliser alors sur cette nouvelle enquête prendront en compte les quotas que vous avez définis.

La méthode par **pondération** ne crée pas une nouvelle enquête, mais une nouvelle variable, nommée « Redressement », qui est ajoutée à la fin de votre questionnaire. Les tableaux que vous exécutez alors dans l'environnement classique (tableaux à plat et croisés) prennent en compte cette pondération. Celle-ci peut se désactiver dans le stade **Traitements et analyses**, menu **Dépouiller**, puis **Options de dépouillement**.

Aide à l'échantillonnage

Depuis le même menu, si votre enquête est un sondage, vous pouvez évaluer la précision de votre échantillon et obtenir l'intervalle de confiance d'un résultat observé. Inversement, vous avez la possibilité de déterminer la taille nécessaire de l'échantillon en fonction de l'erreur acceptée.

9. Traitements et analyse des résultats

Traitements élémentaires : tris à plat et tris croisés

Le traitement et l'analyse des résultats peuvent s'effectuer soit dans l'environnement classique, soit dans l'environnement multimédia : les **tableaux de bord**. C'est cette deuxième série de fonctions qui est aujourd'hui la plus utilisée et c'est celle-ci que nous développerons principalement dans les paragraphes ci-dessous.

Tris à plat et tris croisés dans l'environnement classique

Une première étape permet d'accéder en un clic à l'ensemble des tableaux à plat de l'étude :

Dépouillement automatique

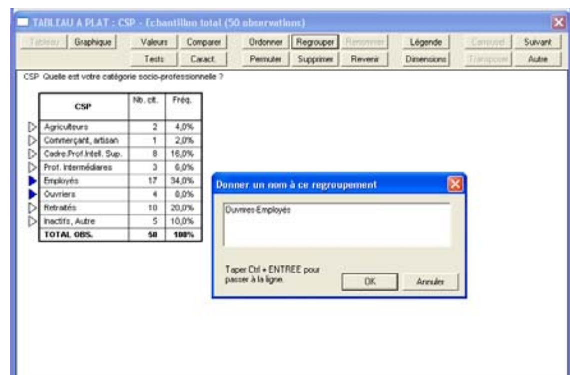
Dans l'environnement classique, cliquer sur **Traitements et analyses** puis sur l'option **Dépouillement automatique** qui donne accès à chacune des variables de l'enquête. Pour aller plus loin dans l'analyse et avoir accès aux options disponibles pour chaque type de variable, on clique toujours sur le bouton **Analyser le tableau** qui se trouve en bas, au centre.



Les fonctions d'analyse du dépouillement automatique

L'environnement d'analyse classique donne accès à différents traitements accessibles sous forme de boutons.

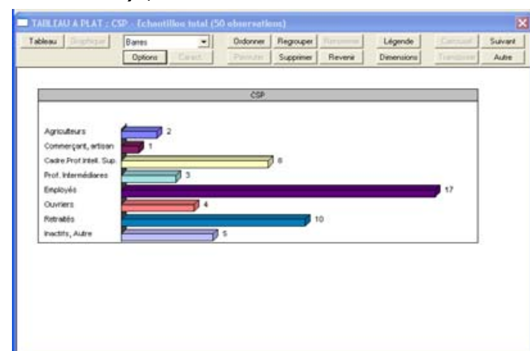
Il permet par exemple de sélectionner une ou plusieurs modalités de la variable étudiée. Pour cela on clique sur la flèche qui se trouve en face de la modalité (qui devient bleue lorsque la modalité est sélectionnée). Pour une sélection multiple, on procédera de même en cliquant successivement sur les différentes modalités à sélectionner, mais tout en maintenant la touche Ctrl du clavier enfoncée.



CSP	Nb. cit.	Fréq.
Agriculteurs	2	4,0%
Commerçant, artisan	1	2,0%
Cadre Prof.Intell. Sup.	8	16,0%
Prof. Intermédiaires	3	6,0%
Employés	17	34,0%
Ouvriers	4	8,0%
Retraités	10	20,0%
Inactifs, Autre	5	10,0%
TOTAL OBS.	50	100%

Les différents boutons permettent :

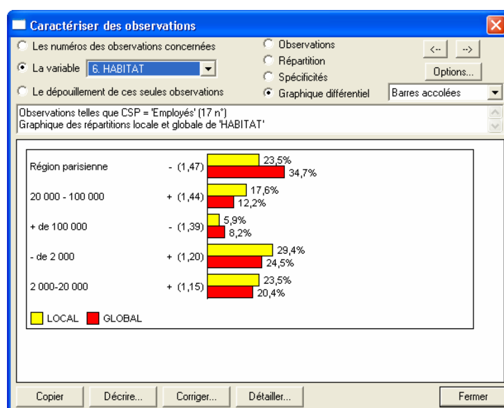
- de passer du tableau au graphique (bouton **Graphique**) ;
- de passer du graphique au tableau (bouton **Tableau**) ;



- de choisir les valeurs à afficher dans le tableau (bouton **Valeur**) ;
- de comparer 2 modalités sélectionnées au préalable (bouton **Comparer**) ;
- de regrouper plusieurs modalités sélectionnées (bouton **Regrouper**) ;
- de renommer une ligne (bouton **Renommer**) ;
- de supprimer une ligne ou une colonne du tableau (bouton **Supprimer**) ;
- d'écrire une légende (bouton **Légende**) ;
- de fixer des dimensions (bouton **Dimensions**) ;
- et enfin de revenir au tableau initial avec le bouton **Revenir** (précisons que les transformations précédemment indiquées n'ont aucun impact sur les données brutes) ;
- le bouton **Suivant** permet de passer à la variable suivante et le bouton **Autre** de choisir une autre variable.

On peut aussi caractériser les individus présents sur une modalité en double-cliquant sur la case correspondante du tableau. On obtient alors :

- les numéros des observations concernées ;
- la répartition des individus sur une autre variable de l'enquête, avec une comparaison avec la population globale (voir **Spécificités** et **Graphique différentiel**) ;
- le dépouillement sommaire de ces individus sur toutes les variables de l'étude.



Tris croisés

Dans le stade **Traitements et analyses**, choisir **Tableaux croisés – Strates** puis le bouton **Tableaux croisés**.

Dans les deux menus déroulants, on choisit la variable en ligne puis la variable en colonne.

On clique alors sur **Croiser V1 et V2** puis sur **Analyser**. Le tableau croisé apparaît alors et l'on peut - si on le souhaite - parfaire sa présentation (regroupements, graphique etc.), selon les options indiquées dans le paragraphe précédent.

Pour vous exercer :

Ouvrir l'enquête Démo dans C:\SphinxV5\Enquêtes

Dans le menu Traitements et analyses, choisir Tableaux croisés – Strates puis le bouton Tableaux croisés.

Choisir la variable n°2 CSP en ligne et la variable n°4 DIPLOME en colonne. Cliquez sur Croiser V1 et V2 puis Analyser.

Sélectionner les deux colonnes Refus et Autres puis cliquez sur l'onglet Supprimer

Sélectionner les deux ligne Agriculteurs et Commerçant, artisan puis cliquez sur l'onglet Regrouper. Nommer ce regroupement Agri/Comm/Artisan par exemple.

Cliquez sur l'onglet Valeurs puis sélectionner l'affichage des pourcentages en ligne.

Cliquez enfin sur l'icône Graphique.

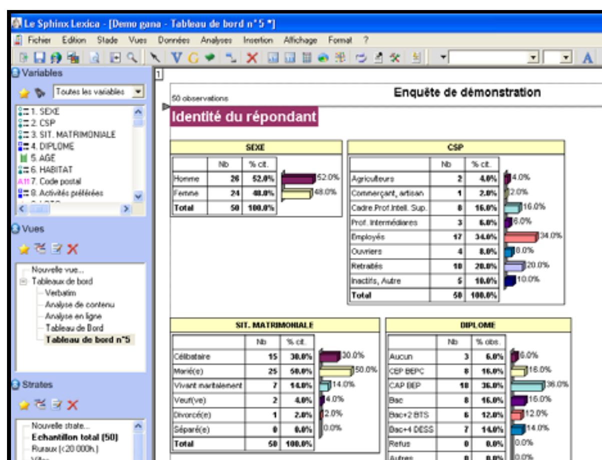


Tris à plat et tris croisés dans l'environnement multimédia : les tableaux de bord

Construction automatique du dépouillement

Dans le menu **Traitements et analyses**, on choisit les **Tableaux de bord**.

La construction automatique est alors disponible. Celle-ci a pré-sélectionné toutes les variables de l'enquête. Vous n'avez alors qu'à choisir un style de présentation puis cliquer sur **Terminer**. On obtient alors le dépouillement de l'ensemble des tableaux à plat et des éventuels groupes de questions de l'enquête.



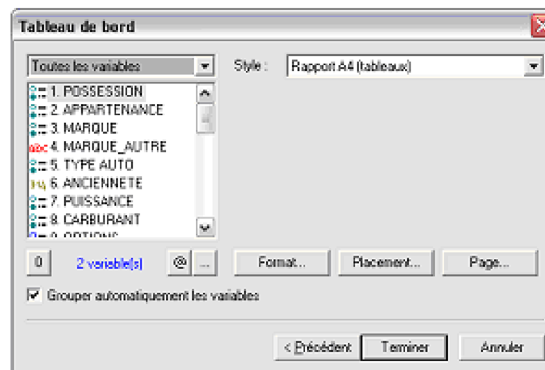
Construction pas-à-pas des tableaux à plat et croisés

Dans la construction pas-à-pas, outre le choix des variables, vous avez également accès à certaines options :

Format : pour paramétrer la présentation des variables (taille, police, couleurs…) dans les tableaux de bord ;

Placement : pour fixer le placement des variables (en ligne, en rond, …) ;

Page : pour définir le format des tableaux de bord (largeur, longueur et couleur de la page, profil attribué, nom…)



Notons que le Sphinx propose également une procédure de présentation automatique de vos tableaux de bord (**Affichage**, **Placer**). Les options de cette rubrique permettent de placer vos objets en colonne, en ligne, en rond, … La taille ainsi que l'emplacement des objets sont modifiables.

Insertion d'un tableau/graphique

On procède exactement comme pour un formulaire multimédia (Voir paragraphe « L'environnement de création du formulaire ») mais ici, les variables apparaissent sous forme de tableaux ou de graphiques…

Pour vous exercer :

En cliquant sur Nouvelle vue au centre à gauche de l'écran, on peut ensuite choisir Tableau de bord puis Construction pas à pas. Cliquez sur 0 puis Terminer pour parvenir à une page vierge.

Choisir la variable n°1 SEXE dans la liste des variables. Maintenir le clic sur la variable et faites la glisser sur la page. On obtient un premier tableau de résultats.

Modification des propriétés de l'objet

En double-cliquant sur un objet (ou clic droit puis **Modifier le format**), on peut en modifier les différentes propriétés, grâce aux différents onglets que nous allons décrire ci-après :

- Contenu : pour indiquer les éléments édités : tableau, graphique, nom ou libellé de variable, description complémentaire, numéro de la question, etc ;

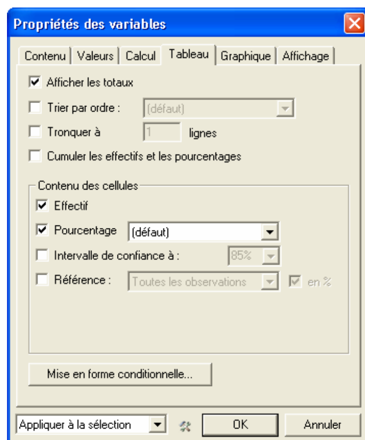
- Valeurs : pour afficher certains indicateurs comme le taux de réponse, la moyenne, la médiane, l'écart-type pour les questions de type numérique ;

- Calcul : pour limiter le tableau à un profil de répondants, ignorer ou non les non-réponses, simplifier un tableau (par ex. suppression ou regroupement) ;

- Tableau : pour indiquer les éléments à afficher dans le tableau et ses éventuelles options de mise en forme (**Mise en forme conditionnelle**) ;

- Graphique : pour choisir son type de graphique, son contenu et son aspect ;

- Affichage : pour modifier les polices de caractère et les couleurs des différents éléments de l'objet.



Pour vous exercer :

Nouvelle vue puis Tableau de bord puis Construction pas à pas. Cliquez sur 0 puis Terminer pour parvenir à une page vierge.

Insérer le tableau de la question n°2 CSP grâce à un cliquer-glisser. Faire un clic droit pour Modifier le format.

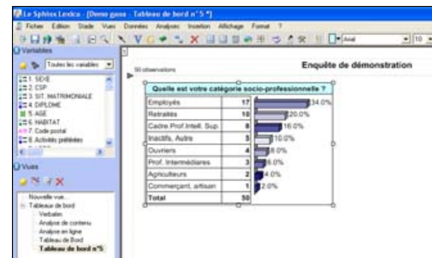
Dans l'onglet Contenu, cochez Graphique mais aussi Libellé et décochez Variable.

Dans l'onglet Valeurs, cochez Taux de réponse.

Dans l'onglet Tableau, cochez Trier…, décochez Pourcentage.

Dans l'onglet Graphique, dans la zone Aspect choisir un dégradé de couleurs à partir d'un bleu foncé (le choix des couleurs est accessible depuis le menu déroulant à côté de la case colorée).

Dans l'onglet Affichage, choisir une couleur de fond bleu clair pour le Nom.



Obtention d'un tableau croisé

Un tri croisé s'obtient de deux manières :

- Choisir le menu **Analyses** en haut de page puis **Tableaux croisés**. Comme dans l'environnement classique, sélectionnez les variables que vous voulez placer en ligne puis en colonne et cliquez sur **OK**.

- Editer sur votre page les deux questions que vous voulez croiser. Sélectionnez-les toutes les deux. Cliquez sur l'icône jaune G ou faites un clic droit puis **Grouper**.

Obtention d’un tableau de groupe

Rappelons qu’un groupe de questions est un ensemble de questions successives dont les modalités de réponses sont identiques. Logiquement, les groupes de questions ont été définis dans le stade de l’**Elaboration du questionnaire** (voir § Question tableau ou groupe de questions). Pour les éditer dans les tableaux de bord :

- dans la liste des variables en haut à gauche de l’écran, un menu déroulant permet d’accéder aux groupes définis dans Sphinx. Au lieu des variables, on a alors la liste des éventuels groupes de l’enquête. Choisir le groupe que l’on veut éditer, faire un cliquer-glisser pour l’insérer sur la vue.

- Choisir le menu **Analyses** puis **Tableaux de groupe de questions.**

Présentation des tableaux de bord sur des strates ou profils

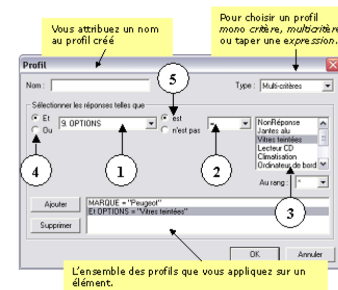
Les strates et profils sont deux techniques utilisées lors de l’analyse des résultats pour vous permettre d’affiner vos analyses en étudiant une sous-population de votre enquête, possédant un ou plusieurs critères que vous aurez défini(s).

Les profils

Un "profil" peut être créé sur un tableau ou un graphique. Pour la création d’un profil :
- faites un clic droit sur le tableau que vous souhaitez modifier,
- sélectionnez l’option **Modifier le format**,
- dans le panneau des propriétés, cliquez sur l’onglet **Calcul**,
- activer l’option **Limiter au profil**.

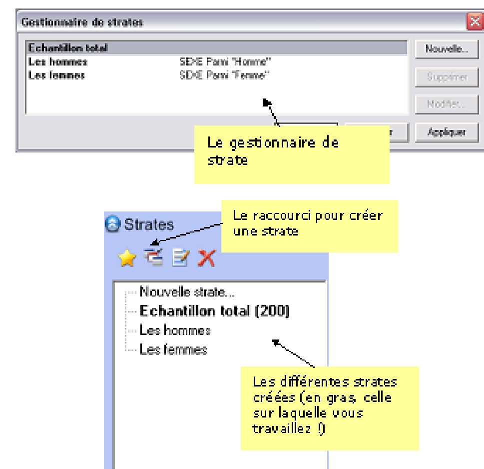
Deux possibilités vous sont proposées. Pour définir un profil "mono-critère", sélectionnez dans la liste la variable puis la modalité à prendre en compte. Pour définir un profil multicritère, vous devez utiliser l’option **Modifier**.

L’écran ci-dessous apparaît alors :



Les strates

Les "strates" se différencient des "profils" par le fait qu’elles s’appliquent automatiquement à l’ensemble des tableaux et graphiques présents sur une vue (le profil peut ne s’appliquer qu’à un ou plusieurs objets d’une page). Pour construire une strate, vous pouvez vous servir du gestionnaire de strates (**Données**, **Strate**) ou sur votre écran utiliser le raccourci prévu à cet effet (voir ci-dessous).



Le gestionnaire de strate offre les options :

- **Nouvelle** permet la création d’une strate,
- **Modifier** pour modifier le profil de la strate,
- **Supprimer** efface définitivement la strate.

Options de calcul pour chaque type de variable

Les paragraphes suivants s'appliquent à des variables spécifiques, pour lesquelles des options particulières sont disponibles.

Variables multiples ordonnées

On peut, pour ce type de variable, ne proposer qu'un seul rang de réponse à l'affichage (par ex. uniquement les modalités citées en 1er rang). On choisit, dans l'onglet **Calcul**, le rang d'affichage.

Variables échelles

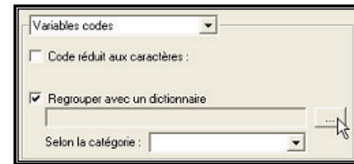
Les variables échelles peuvent être traitées en nombres. Cela signifie que le logiciel associe à chaque échelon une valeur numérique. On peut conserver les valeurs attribuées par défaut, en cochant la case **Traiter les échelles en nombres**. Les valeurs associées vont dans ce cas de 1 à n, n étant le nombre de modalités de la question. On peut également affecter ses propres valeurs numériques à chaque échelon avec **...**

Variables numériques

Les variables numériques peuvent être mises en classes. Le logiciel regroupe dans les mêmes lignes les observations correspondant aux tranches données. Différents types de mises en classes sont proposées : même amplitude, même valeur… Pour réaliser les mises en classes personnalisées, on indique les limites intermédiaires des classes entre points-virgules.

Variables codes

On associe aux variables codes un "dictionnaire" (en format texte) qui traduit chaque code par son contenu. Par exemple, pour un code postal, le dictionnaire traduit pour chaque numéro, le nom du département et la région qui s'y rapportent.



1. Création du dictionnaire : dans un tableur, créer un fichier contenant dans la première colonne les différents codes à traduire (le titre de la colonne est impérativement **Code**). Dans les colonnes suivantes, on intègre les catégories qui définissent chacun des codes (Nommer ces colonnes sans mettre d'espace). Une colonne-une catégorie, une ligne-un code. On enregistre au format texte (séparateur-tabulation) et on ferme le tableur.

2. Sélection des caractères du code : double-cliquer dans la variable et accéder à l'onglet **Calcul**. Cocher la case **Code réduit aux caractères** si l'on souhaite considérer uniquement certains caractères. (Par exemple, pour des codes postaux, on s'intéressera uniquement aux 2 premiers caractères. On écrit donc "1;2")

3. Chargement du dictionnaire : on coche la case **Regrouper avec un dictionnaire** puis clic sur le bouton … Cliquer ensuite sur le bouton d'ouverture de dossier  et sélectionner le fichier texte créé précédemment puis cliquer le bouton **Fermer**.

4. Sélectionner la catégorie à afficher dans le menu déroulant **Selon la catégorie**

Variables dates

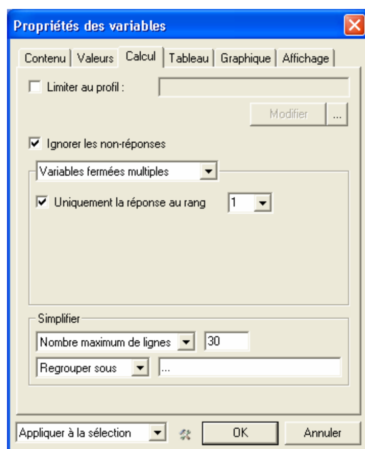
Pour les variables dates, il est possible ici encore d'effectuer différents regroupements par mois, par semaine… On peut aussi effectuer des regroupements personnalisés par période.

Dans ce cas, dans l'onglet **Calcul**, on choisit **En périodes personnalisées** dans le menu déroulant. Puis on écrit les dates séparant les différentes périodes entre points virgules. Par exemple : 01/01/1970 ; 01/01/1990

Simplification et mises en forme conditionnelles

Il est possible d'effectuer différents regroupements pour ajouter du sens à une analyse. L'avantage est que, programmés, ces regroupements, simplifications ou mises en forme conditionnelles seront conservés et mis à jour au fur et à mesure de la saisie.

Environnement de simplification



Double clic dans le tableau à simplifier. Onglet **Calcul**. Dans la partie inférieure du panneau, on accède à l'environnement de simplification.

Limitation du nombre de lignes du tableau

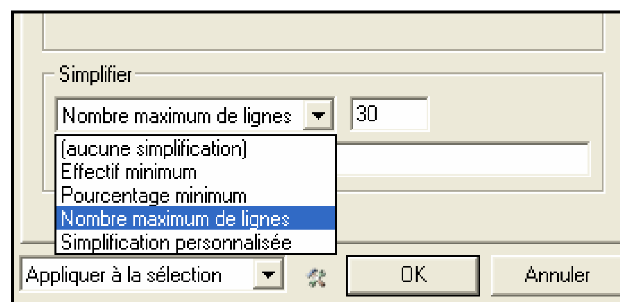
3 options permettent de restreindre le nombre de lignes du tableau :

- **Effectif minimum** : on donne l'effectif minimum à partir duquel une modalité sera affichée dans le tableau

- **Pourcentage minimum** : on donne le pourcentage minimum à partir duquel une modalité sera affichée dans le tableau

- **Nombre maximum de lignes** : le nombre maximum de lignes à afficher dans le tableau.

Pour ces trois cas, on va préciser également ce que l'on souhaite faire des données exclues : soit les regrouper dans une ligne dont on précise le nom, soit les supprimer complètement du tableau. Un menu déroulant est prévu à cet effet.



Supprimer les lignes d'un tableau

Toujours dans l'environnement de simplification, on choisit cette fois ci la **Simplification personnalisée**. On clique sur le bouton ... .

Dans la fenêtre qui s'ouvre alors, on indique les numéros des lignes à supprimer entre parenthèses. Pour supprimer un intervalle de lignes, on les sépare par un tiret. Par exemple, (4-6) supprime les lignes de 4 à 6. Pour supprimer une séquence, on sépare par un point virgule. Par exemple, (4 ;6) supprime les lignes 4 et 6. Si l'on souhaite intégrer plusieurs commandes dans cet environnement, il suffit de revenir à la ligne entre chaque consigne.

Sélectionner des lignes à afficher dans un tableau

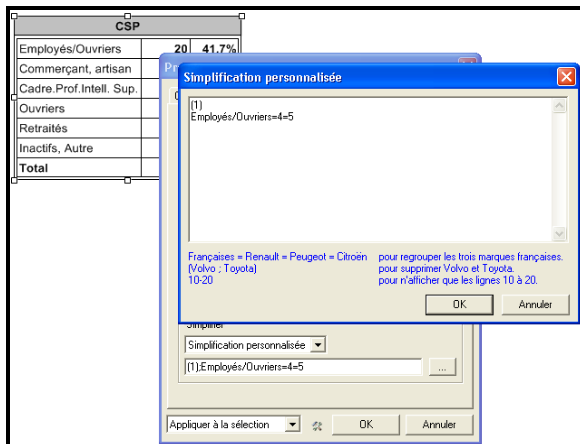
On procède comme précédemment pour la suppression des lignes, mais on enlève les parenthèses. Par exemple, pour n'afficher que les lignes 2 à 8, on écrira simplement : 2-8

Regrouper les lignes d'un tableau

Le regroupement de lignes d'un tableau s'écrit sous la forme d'une séquence de modalités (numéro ou nom de la modalité orthographié avec précision), séparées par des signes égal et précédées du titre du regroupement. Pour regrouper les modalités 1, 2 et 3 sous le nom mongroupe, on écrit : mongroupe=1=2=3. On revient à la ligne pour chaque définition de regroupement.

Cette méthode convient également pour renommer une variable dans un tableau. On indiquera simplement dans ce cas une seule variable ou un seul numéro de ligne en face du titre.

> Pour vous exercer :
>
> Travailler sur l'enquête Démo. Dans les tableaux de bord, éditez le tableau à plat de la variable n°2 CSP.
>
> Double-cliquez sur le tableau pour en modifier les propriétés. Dans l'onglet Calcul et dans la section Simplifier, choisir Simplification personnalisée dans le menu déroulant. Cliquez sur les points de suspension puis saisir :
>
> (1) pour supprimer la première ligne (agriculteurs). Revenir à la ligne puis saisir Employés/Ouvriers=4=5 pour regrouper ces deux catégories.



Les mises en forme conditionnelles

La mise en forme conditionnelle s'active depuis l'onglet **Tableau**, en cliquant sur le bouton **Mise en forme conditionnelle**. Elle permet de paramétrer le contenu des tableaux et graphiques afin de signaler des valeurs remarquables :

- éléments ayant les plus fortes (ou les plus faibles) valeurs,

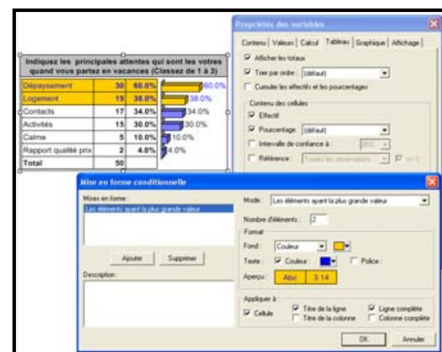
- éléments compris dans un intervalle de valeurs,

- éléments sous ou sur-représentés (en fonction de tests statistiques Chi² ou Student),

- éléments vérifiant une condition plus complexe qui sera alors programmée en basic.

Les valeurs remarquables à mettre en évidence sont choisies dans le menu **Mode**. Le menu **Nombre d'éléments** permet d'indiquer le nombre de valeurs à mettre en évidence. Une (ou plusieurs) condition(s) peuvent être introduites grâce à la liste **Mises en forme** qui permet d'ajouter ou de supprimer des conditions.

Dans la zone **Format**, de nombreuses options sont proposées pour le coloriage des cellules comme le fond de la cellule (couleur ou pictogramme à choisir), la couleur et la police du texte. Depuis la zone **Appliquer à**, le coloriage peut être appliqué au choix, sur le titre ou la ligne/colonne complète du tableau ainsi que sur les éléments graphiques associés (barre, histogramme ou secteur par exemple).



Les tests statistiques

On obtient les résultats des tests statistiques dans l'environnement classique, depuis le bouton **Test** disponible au dessus des tableaux de résultats.

Résultats des tests dans les tableaux de bord

Les tests statistiques s'obtiennent dans les analyses croisées grâce à l'onglet **Tests**. Il suffit de cocher la case **Afficher le test de significativité**. En fonction de la nature des variables qui ont été croisées, le test est affiché : Chi-Deux, F de Fisher ou coefficient de corrélation.

Sont éditées également la valeur de p (le taux d'erreur du test) et entre parenthèses, la significativité du test. TS = très significatif, S= significatif, PS= peu significatif, NS= non significatif. On peut d'ailleurs retrouver ce commentaire complet en clair, en cochant la case correspondante.

L'option **Colorier les cases significatives** permet d'éditer les cellules du tableau où les résultats sont les plus remarquables (Chi-deux partiels les plus élevés par exemple). Ceux-ci apparaissent en bleu (sur-représentations) et en rose (sous-représentations) dans le tableau.

CSP x Diplôme

	Agriculteurs	Commerçant, artisan	Cadre. Prof.Intell. Sup.	Prof. Intermédiares	Employés
Aucun	0	0	0	0	1
CEP BEPC	1	0	0	1	1
CAP BEP	0	0	2	0	11
Bac	1	0	0	1	2
Bac+2 BTS	0	0	2	1	2
Bac+4 DESS	0	1	4	0	0
Total	2	1	8	3	17

p =**0.5%** ; chi2 =**60.38** ; ddl =**35** **(TS)**
La relation est très significative.

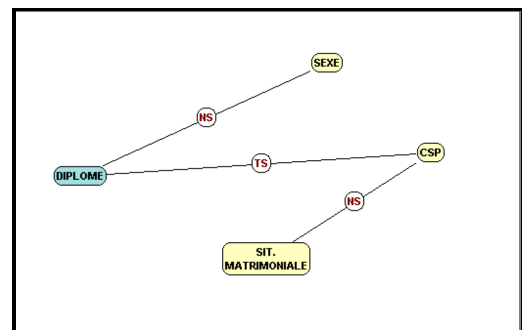
Propriétés des groupes
Contenu | Tests | Calcul | Tableau | Graphique | Affichage
Colorier les cases significatives
Afficher le test de significativité
Commentaire
Appliquer à la sélection | OK | Annuler

Les graphes de relations (*)

Les graphes de relation permettent de tester simultanément les relations statistiques parmi un ensemble de variables donné. Depuis les tableaux de bord, on demande une nouvelle vue puis on choisit **Graphes de relation**.

Formuler des hypothèses (*)

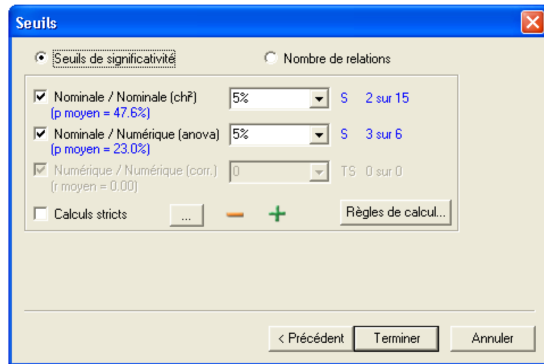
Le bouton **Formuler des hypothèses** est sélectionné. On choisit alors la liste des variables que l'on veut étudier, puis on clique sur **Terminer**. Les variables apparaissent alors sur la vue, sous la forme de "pastilles". Grâce à l'icône correspondant , on va tracer la relation entre les variables à étudier et obtenir immédiatement la significativité du test.



Trouver les principales relations (*)

Le bouton **Trouver les principales relations** est sélectionné. On choisit les variables à étudier.

Un menu apparaît alors qui permet éventuellement de jouer sur les seuils de significativité, afin d'obtenir un nombre plus ou moins grand de relations en faisant varier la sévérité des seuils. Les menus déroulants permettent de faire varier les seuils, les boutons + et – permettent d'obtenir plus ou moins de relations.



On obtient alors un graphe de relations, où seules les relations significatives sont affichées (selon les choix faits au préalable).

Expliquer la variable (*)

Cette option permet de croiser simultanément une variable-clé de l'enquête (par exemple une note de satisfaction globale) avec toutes les autres, afin de déterminer quelles sont les variables qui l'influencent le plus fortement.

Propriétés de la relation (*)

En réalisant un clic droit puis **Modifier le format**, on peut modifier les caractéristiques affichées. On peut ainsi faire apparaître la valeur du test statistique correspondant ou éditer les principaux éléments : chi-deux partiels les plus forts ou moyennes les plus différentes par exemple. **Tester la relation** colorie le trait en fonction de sa significativité.

A partir du graphe de relations, on peut faire apparaître le tableau croisé correspondant, grâce à un clic droit sur la flèche puis **Analyser**. On peut ainsi visualiser les données initiales et les raisons pour lesquelles la relation s'avère significative ou non.

Pour vous exercer :

Utilisez l'enquête Demo.

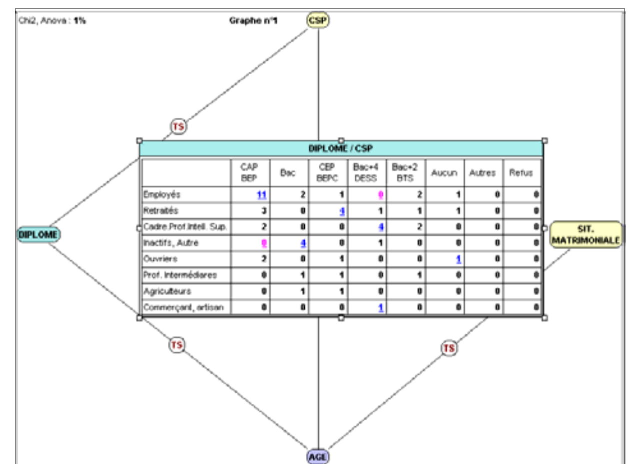
Dans l'environnement des tableaux de bord, choisir Nouvelle vue puis Graphes de relations.

Dans la liste des variables, sélectionnez les questions n°1 à 8. Cliquer sur Trouver les principales relations. Baissez le seuil de significativité à 1% en cliquant sur l'icône . Cliquez sur Terminer.

Répondez Non à la question posée et les variables indépendantes ne seront pas présentes sur le graphe.

Le graphe vous est alors proposé, qui nous permet de constater que les relations suivantes sont très significatives : Diplôme x CSP, Diplôme x Age, Age x CSP et Age x Situation matrimoniale.

Faire un clic droit sur la flèche entre Diplôme et CSP puis choisir Analyser, le tableau croisé entre les deux variables apparaît alors.



DIPLOME / CSP

	CAP BEP	Bac	CEP BEPC	Bac+4 DESS	Bac+2 BTS	Aucun	Autres	Refus
Employés	11	2	1	0	2	1	0	0
Retraités	3	0	4	1	1	1	0	0
Cadre Prof.Intell. Sup.	2	0	0	4	2	0	0	0
Inactifs, Autre	0	4	0	1	0	0	0	0
Ouvriers	2	0	1	0	0	1	0	0
Prof. Intermédiares	0	1	1	0	1	0	0	0
Agriculteurs	0	1	1	0	0	0	0	0
Commerçant, artisan	0	0	0	1	0	0	0	0

Les analyses multi-variées (*)

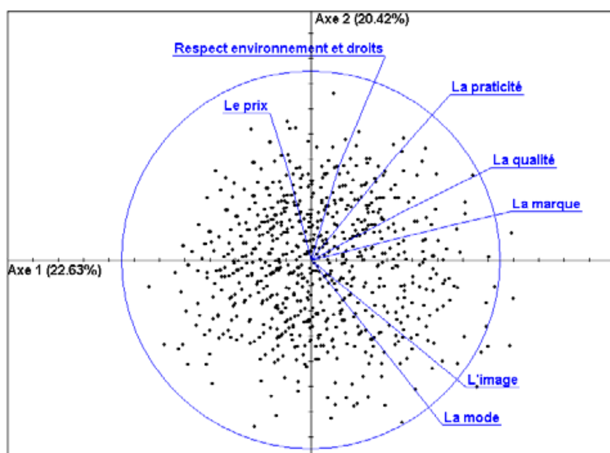
A partir de sa version Plus², le Sphinx permet d'accéder à plusieurs fonctions d'analyse de données multi-variée. Celles-ci sont disponibles dans l'environnement classique mais aussi dans les tableaux de bord.

A.C.P. et A.F.C. (*)

Dans **Traitements et Analyses**, le menu **Approfondir** permet d'accéder aux analyses factorielles : **analyse factorielle multiple** pour l'A.F.C. et **analyse en composantes principales** pour l'A.C.P.

Dans les menus correspondants, il s'agit tout d'abord de sélectionner les variables à analyser, celles-ci doivent être numériques pour l'ACP et nominales pour l'AFC.

On accède alors directement à la carte factorielle où les données sont représentées sur les deux premiers axes.



On peut ensuite accéder à plusieurs fonctions pour aller plus loin dans les analyses. De nombreux résultats numériques sont disponibles : matrice des corrélations, coordonnées des variables ou des modalités, contributions aux axes, etc.

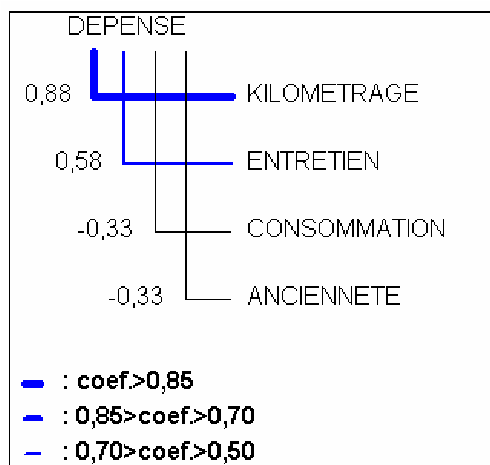
Il est également possible de modifier la **Présentation** de la carte et d'enregistrer sous forme de nouvelles **Variables** les coordonnées de chaque observation sur les axes.
Un bouton permet d'activer une **Typologie** interactive, détaillée à la page suivante.

Régression multiple (*)

Le menu **Approfondir** propose un menu de **Régression Multiple**. Celle-ci peut-être menée globalement ou selon une méthode **Pas-à-pas** ascendante. Les variables à étudier sont forcément numériques.

Sont fournies alors :

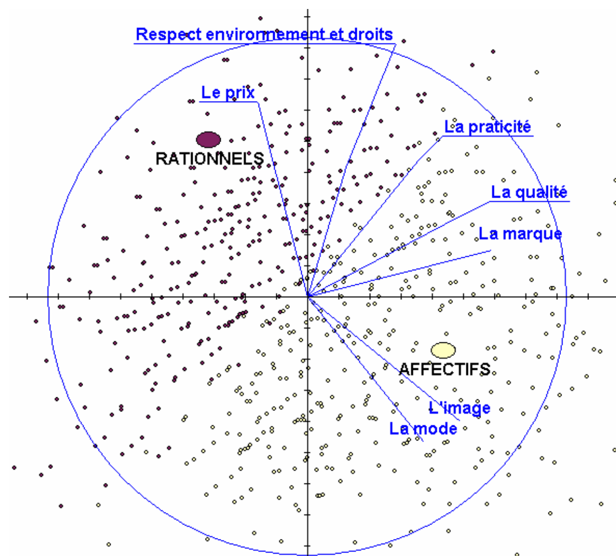
- l'équation de la droite de régression multiple,
- le coefficient de corrélation multiple (r),
- la matrice des corrélations,
- les indicateurs de significativité des paramètres du modèle,
- des représentations visuelles en diagrammes (voir ci-dessous).



Analyse typologique (*)

Le Sphinx propose deux méthodes de construction de typologies : l'une est interactive, l'autre est automatique.

La première s'active depuis une analyse factorielle par le bouton **Typologie**. Elle permet de constituer des groupes de répondants en positionnant soi-même les centres de classes sur la carte. On peut ensuite enregistrer l'appartenance à la classe dans une nouvelle variable fermée (bouton **Variables**).



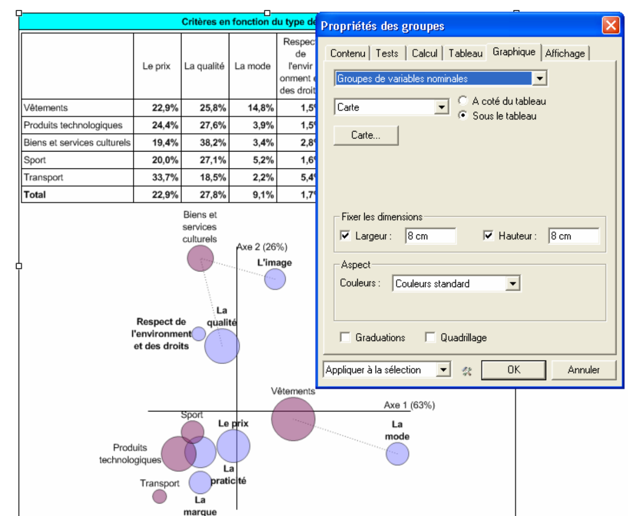
La **Classification automatique** (selon la méthode des centres mobiles) s'active depuis le menu **Approfondir**. Une détermination automatique est disponible, ou une fonction de construction imposée de n classes. On peut également demander une partition initiale sur une des variables de l'enquête.

Sont ensuite fournies des statistiques permettant d'évaluer chaque typologie proposée. On peut enregistrer la classe dans une nouvelle variable fermée.

Les analyses factorielles dans les tableaux de bord (*)

Vous pouvez réaliser vos analyses multi-variées dans les tableaux de bord et bénéficier ainsi de tous leurs attraits : mise à jour selon la saisie ou la strate choisie, enregistrement du format, analyse en ligne etc.

Dans les tableaux de bord, les cartes factorielles sont un des "graphiques" que l'on peut choisir afin d'illustrer un tableau représentant plusieurs variables. On peut les obtenir par exemple depuis un groupe de questions ou un tableau croisé. Dans le menu déroulant **Graphique**, choisir **Carte**. On voit alors la carte factorielle où les données sont représentées sur les deux premiers axes.



Le bouton **Carte** permet d'accéder à des fonctions plus avancées de représentation et d'accès aux données numériques comme les diverses coordonnées et contributions.

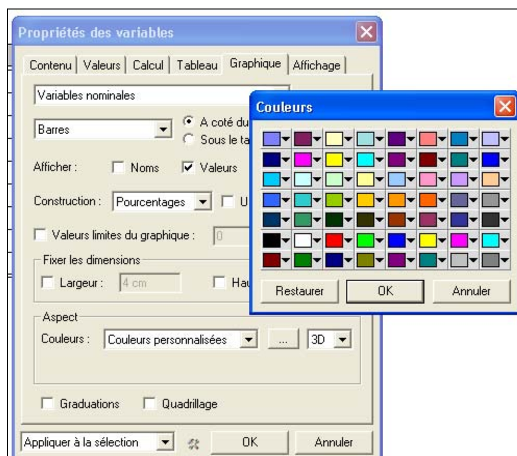
Les graphiques

Types de graphiques

Dans les propriétés d'une variable, l'onglet **Graphique** permet de faire varier la représentation graphique des résultats. Les principaux types de graphiques sont disponibles dans un menu déroulant : secteur, histogramme, barre, courbe, aire, anneau. Des graduations et un quadrillage peuvent être édités, les valeurs limites du graphique peuvent être modifiées grâce aux options correspondantes.

Choix des couleurs

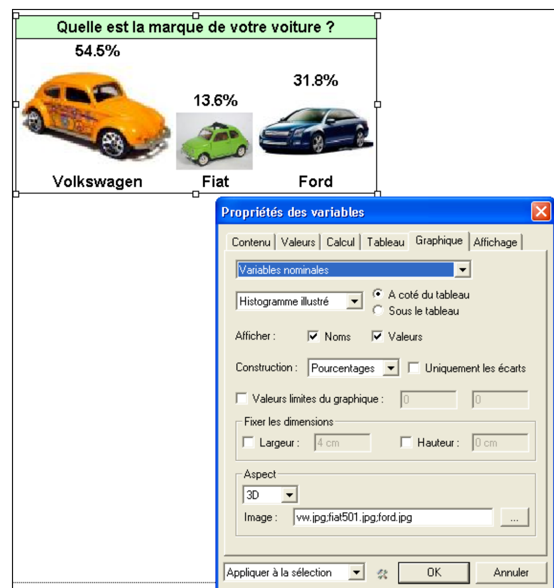
Par défaut, les couleurs proposées sont les couleurs standard du logiciel mais celles-ci peuvent être modifiées dans un menu déroulant. On peut choisir une seule couleur ou un dégradé de couleur. Les **Couleurs personnalisées** (après un clic sur l'icône ...) donnent accès à une palette qui permet d'affecter une couleur déterminée pour chaque modalité du tableau. La palette se parcourt de haut en bas et de gauche à droite.



La première case (bleue par défaut) est affectée aux non-réponses, la deuxième (grenat) à la première modalité, la troisième (jaune) à la deuxième modalité et ainsi de suite… C'est en utilisant la liste déroulante présente à côté de chaque couleur qu'on pourra ainsi la modifier.

Histogramme illustré

Depuis l'onglet **Graphique**, il est possible d'associer à un graphique en histogramme n'importe quelle illustration, fournie dans le logiciel ou récupérée par ailleurs. Les barres de l'histogramme se dimensionnent alors en fonction des fréquences des modalités. On peut associer soit une seule et même illustration pour toutes les barres, soit une illustration spécifique pour chaque modalité. Les coordonnées des illustrations à afficher sont indiquées dans l'espace correspondant, s'il y a plusieurs illustrations, leurs noms sont séparés par des points-virgules.



Cartes géographiques

Insérer une carte par l'article **Carte géographique** du menu **Insertion.** On accède alors au dialogue de paramétrage de la carte.

On choisit la carte dans la liste déroulante des cartes disponibles. En standard, les cartes suivantes sont proposées : la carte de France par régions, par départements (avec ou sans les départements d'outremer), les cartes de chaque région par département et un exemple de carte d'un département (Var) par communes (ou plus exactement par code postal).

On indique quel est l'**Identificateur**, c'est à dire la variable de l'enquête qui contient l'information géographique, par exemple le Code Postal.

Le **Calcul** le plus courant consiste à afficher le nombre d'observations, mais on peut aussi afficher une moyenne (la moyenne d'âge par ex.), une somme (la surface cumulée d'une série de magasins par ex.) ou une autre formule de calcul.

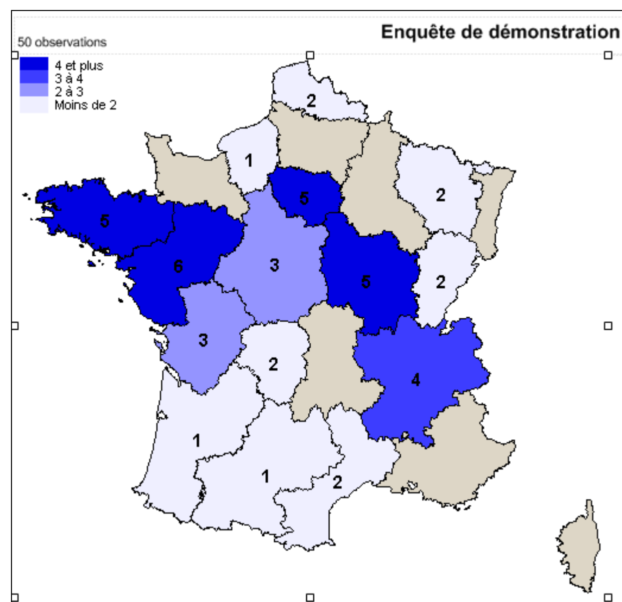
On choisit pour cette carte des options d'**Affichage** : la couleur des contours, la couleur du fond, l'affichage optionnel des noms des objets de la carte.

Le Sphinx va proposer une **Présentation des données** sous formes de valeurs, de symboles ou de remplissage.

Dans le dialogue de propriétés, on trouvera les éléments habituels pour ajouter un titre, paramétrer la police de ce titre, limiter le calcul à un profil d'observations, etc.

Pour vous exercer :

Dans l'enquête Démo, depuis les tableaux de bord, choisir le menu Insertion puis Cartes géographiques. Choisir une représentation par régions. L'identificateur est ici le Code Postal. Montez la police de caractères à 16. Changez la couleur de remplissage en dégradé de bleu. Cliquez sur OK.



Re-codification et enrichissement de données

Le Sphinx permet de créer de nouvelles variables, par transformation, modification ou calcul sur les variables initiales de l'enquête. Les sections suivantes présentent une partie des fonctions disponibles.

Recoder une variable numérique

Une variable définie dans le questionnaire comme ouverte numérique peut être mise en classes ultérieurement afin de créer une variable fermée, dont les modalités seront les classes définies par le chargé d'études.

Cette opération se fait depuis l'environnement classique du Sphinx. Dans **Traitements et analyses**, choisir le menu **Recoder** puis **Mettre les numériques en classes**. Choisir la variable à recoder dans le menu déroulant. Deux méthodes sont alors disponibles :

- la première consiste à placer les limites de classes manuellement, grâce aux petits triangles que l'on peut faire glisser. On peut supprimer ou ajouter des bornes grâce aux boutons correspondants.

- la seconde revient à indiquer les limites de classes, depuis le bouton **Editer**, séparées par des points-virgules.

Une fois l'opération terminée, on va créer la nouvelle variable en cliquant sur **Recoder sur une échelle**.

Transformer une variable (*)

Depuis le menu **Recoder**, cette fonction permet notamment de changer le type d'une variable, de décomposer une variable multiple ou de regrouper définitivement les modalités d'une variable fermée. Seule cette dernière fonctionnalité est présente dans la version Primo.

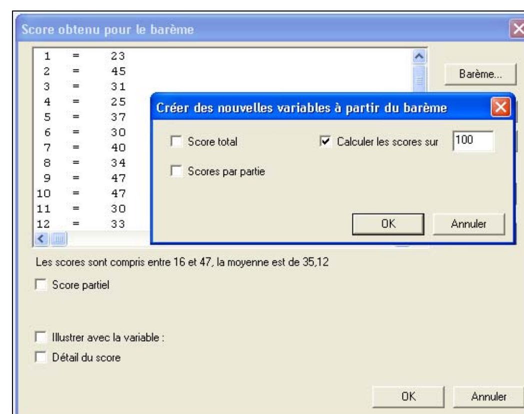
Barèmes

Dans l'**Elaboration du questionnaire**, on peut associer des valeurs numériques aux différentes modalités de réponses à une question.

Ultérieurement, dans **Traitements et analyses**, on peut créer une nouvelle variable sur la base du score obtenu avec le barème. Voir le menu **Dépouiller**, **Appliquer un barème** puis valider avec **OK**.

Dans cette section, on peut illustrer le score avec une autre variable de l'enquête, produire un score partiel (sur une partie seulement des questions) ou obtenir le détail du barème.

Le bouton **Recoder** permet de créer une nouvelle variable numérique avec le score obtenu.



Pour vous exercer :

Utiliser l'enquête Démo. Un groupe de questions échelles a été créé. Vérifier qu'un barème a été appliqué à chacune des 5 questions du groupe (le rapport qualité-prix ; le dépaysement ; le logement ; les activités ; le contact avec l'habitant). Le barème est le suivant, on a attribué 10 points pour très satisfait, 7 pour satisfait, 5 pour indifférent, 3 pour insatisfait et 0 pour très insatisfait.

Passer au stade Traitements et analyses puis dans le menu Dépouiller et Appliquer un barème.

On obtient alors le score de chaque répondant. Cliquer sur Recoder puis on coche Calculer les scores sur 100. Saisir un nom de variable (satisfaction sur 100 par ex.) puis cliquer sur OK. On aura alors créé un score de satisfaction total sur 100 pour chaque répondant.

Calculer une nouvelle variable (*)

Dans le menu **Traitements et analyses**, puis **Recoder** et **Calculer une nouvelle variable**, on a accès à des fonctions de calcul simples ou sophistiquées pour créer de nouvelles variables.

Les opérations mathématiques classiques sont accessibles, mais aussi des formules ou procédures plus complexes à programmer en Basic.

Analyse des données qualitatives

Ouverture d'un texte (**)

Le Sphinx (*seulement dans sa version Lexica*) permet d'ouvrir des textes (entretiens, tables rondes, articles, bases bibliographiques etc.) saisis au préalable dans un traitement de textes. Il est indispensable que le texte soit saisi selon un format bien précis, avec des repères réguliers pour les différentes parties du texte, nommées "balises" ou "annotations". Il sera enregistré en format texte (.txt).

A titre d'exemple, le texte d'entretiens serait saisi de la manière suivante :

Entretien : Paul

Sexe : Homme

Age : 37

CSP : Employé

Q : Qu'est-ce que ça signifie pour toi le plaisir de manger ?

R : Ha, j'adore, parce que d'abord, j'aime bien faire la cuisine, et puis je ne sais pas, tu as du plaisir, de préparer une bonne table, c'est beau à voir, ce que tu as fait, t'apprécies.

Entretien : Annie

Sexe : Femme

Age : 18

CSP : Etudiante

Q : Je voudrais qu'on parle du plaisir de manger, du plaisir que tu prends à manger.

R : Ce que j'adore manger, c'est la viande, rouge, saignante. Pas de sucre, déjà, j'ai horreur du sucre, pas de gâteaux. Etc.

L'ouverture du texte se fait depuis le menu de démarrage du Sphinx, par la fonction **Texte à analyser**. On indique l'emplacement du fichier texte à importer, puis on coche si les textes sont annotés ou balisés.

Pour les textes balisés, dans l'écran suivant, on a la possibilité d'indiquer les caractères qui constituent les balises ; par défaut, il s'agit du retour à la ligne (^p) et des deux points (:). L'écran suivant n'est utile que pour un paramétrage plus pointu, dans le cas d'un texte mal structuré. Par défaut, on clique sur **OK**.

On peut ensuite valider la reconnaissance des différentes balises et de préciser le cas échéant le type des différentes variables que l'on va créer dans le Sphinx. En cliquant une dernière fois sur **OK**, une "enquête" au format Sphinx est créée.

Les verbatims

La fonction **Verbatim** consiste à restituer le texte des réponses, dans son intégralité ou en opérant des sélections sur les informations souhaitées ou sur un profil d'individus par exemple.

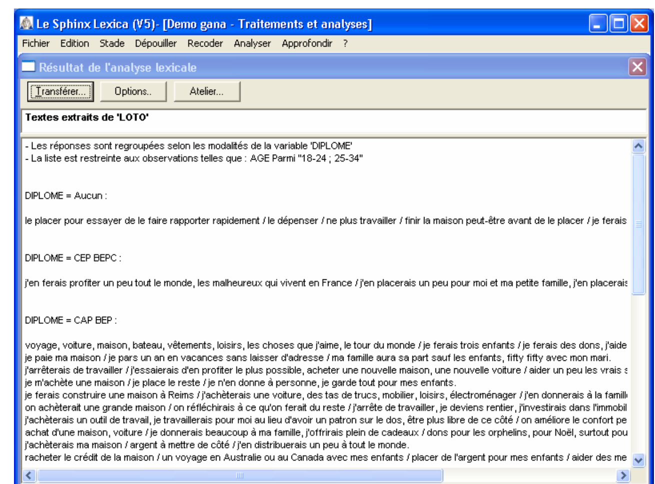
On y accède depuis les **Traitements et analyses**, dans le menu **Recoder** puis **Etudier les textes**. Choisir la variable à étudier puis cliquer sur **Verbatim**. On a alors la possibilité de sélectionner les textes à extraire et :

- de les trier en fonction des modalités d'une question,
- de limiter l'extraction à un profil d'observations,
- de se concentrer sur les extraits contenant un ou plusieurs mots.

On peut ensuite imprimer ou transférer les verbatims vers un rapport (bouton **Transférer**).

Pour vous exercer :

Ouvrez l'enquête Démo. Depuis traitements et analyses, choisissez Recoder puis Etudier les textes. Choisissez la variable n°9 LOTO et cliquez sur Verbatim. Cochez "Trier selon" et choisissez la variable n°4 DIPLOME. Cochez "Réduire à un profil". Dans le menu correspondant, choisissez la variable AGE et sélectionnez les tranches 18-24 et 25-34. Cliquez deux fois sur OK. La liste des réponses sélectionnées est alors éditée. Vous pouvez l'imprimer ou l'exporter dans un autre logiciel en cliquant sur Transférer.



L'analyse de contenu

Cette méthode consiste ici à recoder une variable texte en une ou plusieurs variables fermées, en lisant toutes les réponses et en les affectant à une ou plusieurs catégories thématiques. Ces catégories ou thèmes sont pré-définis (code-book) ou créés au fur et à mesure de la lecture.

L'analyse de contenu est adaptée à un corpus de petite ou moyenne taille (moins de 500 observations par ex.)

Analyse de contenu en cours de lecture

Celle-ci peut s'effectuer dans l'environnement classique du Sphinx.

Depuis les **Traitements et analyses**, choisir le menu **Recoder** puis **Etudier les textes**. Sélectionner la variable à analyser puis cliquer sur **Analyse de contenu**.

S'il s'agit d'une nouvelle codification en une variable fermée (cas le plus fréquent), on définit au préalable le nombre de modalités de réponses et le nombre de réponses possibles. L'étape suivante consiste à nommer la nouvelle variable. On arrive alors dans l'environnement d'analyse de contenu.

En haut de l'écran, on peut lire chaque réponse. La navigation se fait grâce aux flèches en bas à gauche. Le bouton **Changer les modalités** permet de définir, modifier ou ajouter des thèmes. L'analyse de contenu consiste ensuite à cocher le (ou les) thème(s) correspondant à la réponse lue. Une nouvelle variable fermée s'incrémente alors, au fur et à mesure du codage et pourra être analysée comme les autres variables de l'enquête.

Analyse de contenu avec code-book

Celle-ci doit s'effectuer dans l'environnement multimédia du Sphinx. Il s'agit de créer plusieurs questions pour recoder les réponses sur plusieurs dimensions : le thème, le sous-thème, la tonalité, les acteurs évoqués par exemple… On peut également prévoir une nouvelle question texte pour y recopier les "perles" ou les réponses particulièrement intéressantes.

On commence par la création de nouvelles questions dans l'élaboration du questionnaire. On prépare ensuite un environnement personnalisé d'analyse de contenu dans les formulaires multimédia. L'analyse de contenu s'effectue depuis le menu **Consulter/Modifier** de l'environnement multimédia.

Pour vous exercer :

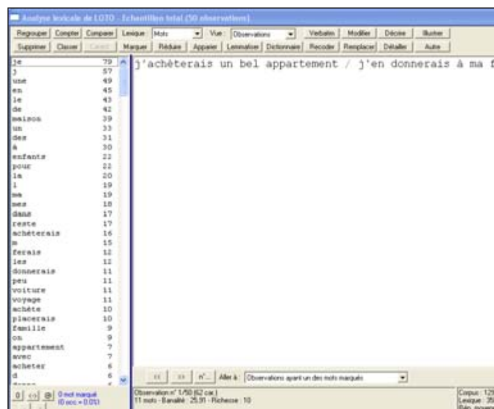
Ouvrez l'enquête Démo. Des questions ont déjà été créées pour l'analyse de contenu dans la partie Grille pour l'analyse de contenu.

Dans les formulaires multimédia, créez un nouveau formulaire écran. Cliquez sur 0 sous la liste des variables puis sur Terminer. Selon le modèle ci-dessus, placez alors dans le formulaire : le texte de départ (variable n°9 LOTO) et les variables d'analyse de contenu (n°22 à n°31). Placez-vous dans le stade Consulter-Modifier pour tester la codification.

L’analyse lexicale (*)

Elle est adaptée à un corpus de moyenne ou de grande taille (au moins 200 observations par ex.). En deçà, l’analyse de contenu est plus précise. L’analyse lexicale consiste à appréhender le texte par son lexique, soit la liste complète des mots et/ou expressions cités. A partir de ce lexique, on peut :

- identifier les mots et expressions les plus cités ;
- constituer des regroupements thématiques ;
- naviguer rapidement dans le texte, selon des critères de recherche ;
- créer des mesures lexicales ;
- recoder le texte en nouvelles variables fermées etc.



Pour réaliser ces opérations, le Sphinx (dans ses versions Plus² et Lexica) donne accès à un atelier lexical, présentant :

- à gauche, le lexique classé par défaut par nombre décroissant d’occurrences,

- à droite, le texte à une réponse,

- en haut, une liste de fonctions,

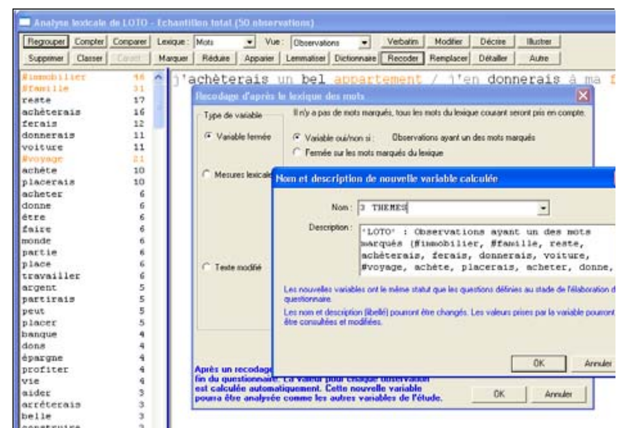
- en bas, des flèches pour naviguer parmi les réponses, selon des critères définis dans le menu déroulant.

Pour vous exercer :

Ouvrez l’enquête Démo. Depuis Traitements et analyses, choisissez Recoder puis Etudier les textes . Choisissez la variable n°9 LOTO et cliquez sur Atelier lexical.

Vous obtenez à gauche le lexique brut. Cliquez sur le bouton Réduire puis sur Ignorer les éléments du dictionnaire des mots-outils. Si celui-ci n’est pas actif, cliquez sur Dictionnaires et choisir le fichier situé en C:\SphinxV5\Dictionnaires\Mots-outils.dic. Cliquez sur OK.

Sélectionnez ensuite dans le lexique les mots maison et appartement, cliquez sur Regrouper, nommez le regroupement immobilier. Sélectionnez enfants et famille, regroupez-les sous famille. Sélectionnez voyage, voyages et vacances, regroupez-les sous voyage. Trois groupes sont alors constitués, leurs noms apparaissent désormais en orange dans le lexique, précédés d’un #. Sélectionnez ces groupes dans le lexique. Cliquez sur Recoder puis sur Variable fermée / Fermée sur les mots marqués du lexique. Cliquez sur OK. Donnez un nom à cette nouvelle variable "3 Thèmes" par exemple. Cliquez sur OK. Dès lors, cette nouvelle variable fermée pourra être analysée comme les autres variables de l’enquête.



Communication avec d'autres applications

Le Sphinx est un logiciel très ouvert, tant en amont pour l'importation de données existantes dans d'autres applications (de type tableur par ex.), qu'en aval pour le transfert vers d'autres programmes (de type analyse statistique par ex.).

Récupération des données existantes (*)

L'ouverture de données pré-existantes se fait depuis le menu de démarrage de Sphinx, dans la fonction **Données externes**. On choisit le type de fichier à importer et sa provenance (fichier de type Excel®, Access® ou étude créée dans SPSS® par exemple). On indique ensuite la localisation du fichier concerné puis on clique sur **Ouvrir**.

On accède alors à des fonctions d'importation en plusieurs étapes :

- Visualisation des données repérées et éventuellement changement de feuille dans le menu déroulant. Cliquer sur **Suivant** ;
- Accès à quelques options d'importation : ignorer les n premières lignes ou modification des séparateurs. Cliquer sur **Suivant** ;
- Détection automatique du type des variables. Possibilité de modification du nom et du type des variables. Cliquer sur **Suivant** ;
- Détection d'éventuelles erreurs. Cliquer sur **Terminer.**

Une nouvelle enquête Sphinx va se créer avec le nom que vous aurez saisi. Vous pouvez désormais appliquer tous les traitements disponibles dans Sphinx sur cette base de données importée.

Pour vous exercer :

Depuis n'importe quelle enquête Sphinx, revenez sur l'écran de démarrage puis choisissez Données externes. Cliquez sur Un fichier de données puis sur OK. Allez chercher un fichier d'exemple dans c:\SphinxV5\Enquêtes\Exemples\Ouvrir un fichier de données\Wines-excel.xls. Aux étapes 1 et 2, cliquez sur Suivant. A l'étape 3, vous pouvez changer le nom des variables en français : année, couleur, garde et prix. Changez le type de la variable année de numérique à fermée unique. Cliquez sur Suivant. A l'étape 4, cliquez sur Terminer. Nommez l'enquête vins.que et enregistrez-la.

Exportations vers d'autres applications

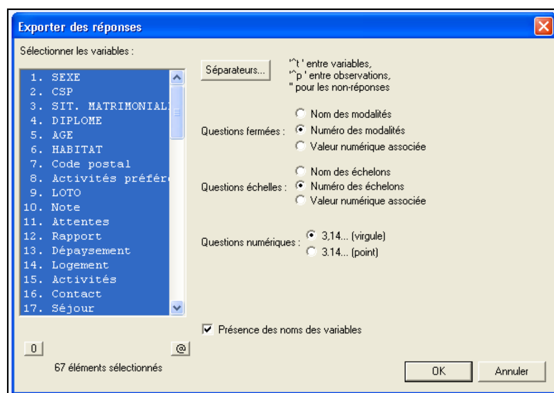
Les données saisies (données brutes) dans le Sphinx peuvent être exportées vers d'autres applications : tableurs, bases de données ou logiciels statistiques. L'opération se fait depuis le menu **Fichier**. Choisir **Faire une sauvegarde** / **Exporter**.

Exporter vers un tableur

On coche alors l'option **Exporter les données**. On accède à un menu qui permet de sélectionner les variables à exporter (surlignées en bleu). On peut préciser quelques options : nature des séparateurs, numéro des modalités ou modalités en clair, point ou virgule pour les nombres décimaux, présence ou non du nom des variables. Cliquer sur OK. Choisir le nom et l'emplacement du fichier de données que vous avez généré.

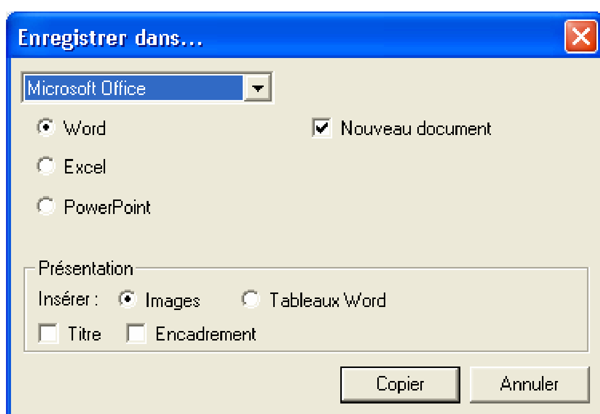
Convertir l'enquête

Depuis **Faire une sauvegarde** / **Exporter**, on coche l'option **Convertir l'enquête pour**. On a ici notamment la possibilité d'enregistrer le fichier au format SPSS® (.sav), triple-s ou dans le format d'une version antérieure du Sphinx.



Communication avec Microsoft Office® et Open Office®

Dans les menus de l'environnement multimédia (formulaires ou tableaux de bord), le menu **Fichier** puis **Enregistrer dans Office** permet de créer directement des fichiers pour Microsoft Office® (Word®, Excel®, Powerpoint®) ou pour Open Office® contenant les pages de vos formulaires ou de vos tableaux de résultats. Après avoir sélectionné l'application souhaitée, cliquer sur **Copier** pour créer les documents correspondants. Pour Word®, on peut sélectionner une exportation des tableaux en **Images** ou en **Tableaux Word** modifiables.



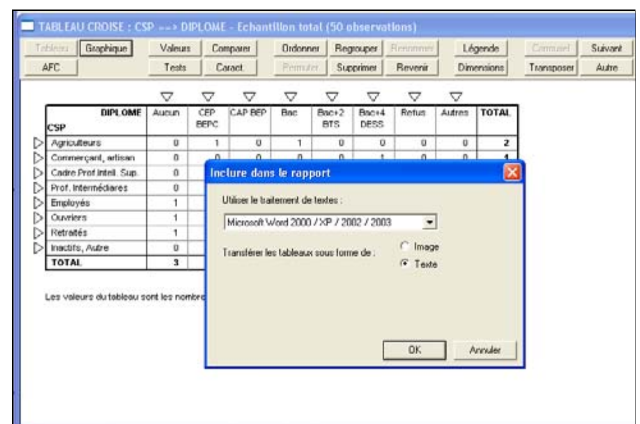
La présentation des rapports

Rapports papier

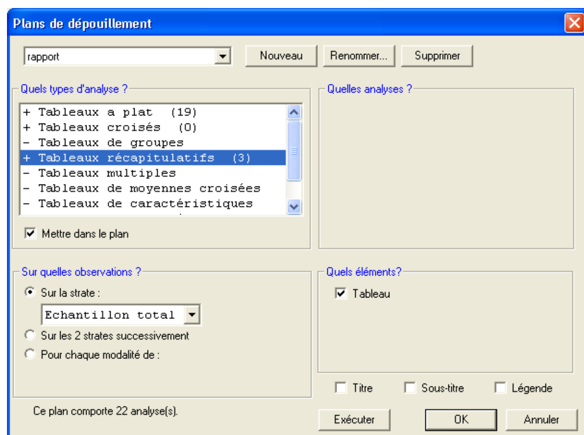
On peut construire aisément des rapports d'étude dans un traitement de textes, selon plusieurs méthodes possibles :

La première consiste à créer un fichier dans le traitement de textes pour y insérer tableaux ou graphiques au fur et à mesure, grâce tout simplement à des copier/coller.

La deuxième, dans l'environnement classique du Sphinx, utilise la fonction **Inclure dans le rapport** dans le menu **Fichier** (raccourci Ctrl+W). Celle-ci permet d'envoyer directement le tableau ou le graphique dans Word®, en image ou en tableau.



La troisième, toujours dans l'environnement classique du Sphinx, consiste à créer un **Plan de dépouillement** (depuis **Traitements et analyses**, menu **Dépouiller**). Dans cet environnement, on sélectionne successivement les différentes analyses que l'on souhaite éditer, les éléments à transférer (tableau et/ou graphique) et éventuellement la population sur laquelle on veut obtenir les résultats.



Pour vous exercer :

Ouvrez l'enquête Démo. Dans Traitements et analyses, choisissez le menu Dépouiller puis Plan de dépouillement.

Sélectionnez la ligne Tableaux à plat, dans la liste à droite, cliquez sur les variables 1, 2 et 3 (en maintenant la touche Shift enfoncée). Cliquez sur tableaux et graphiques dans la partie Quels éléments ?. Dans les tableaux croisés, sélectionner le tri normalement disponible (CSP x DIPLOME). Décochez la case Mettre dans le plan pour la ligne des Tableaux récapitulatifs.

Cliquez sur Exécuter puis choisissez de transférer les tableaux en texte. Nommez votre fichier Word® et cliquez sur Enregistrer. Les éléments souhaités sont transférés dans Word®.

Rapports depuis les tableaux de bord

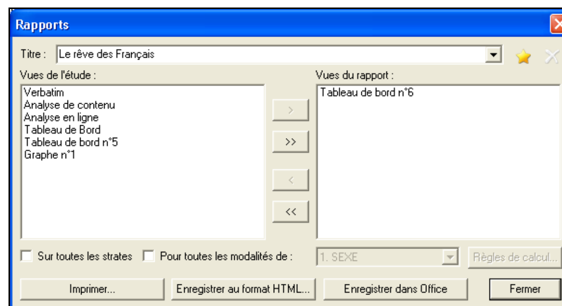
Dans l'environnement multimédia, rappelons que les tableaux de bord permettent en fait de créer des pages dans un contexte semblable à celui d'un traitement de textes. On insère des tableaux, des graphiques et des légendes.

On peut ensuite soit :

- imprimer directement les tableaux de bord,

- **Préparer un rapport** dans le menu **Fichier**. Cette fonction permet de sélectionner une ou plusieurs vues, de les faire porter sur une ou plusieurs strates ou sur toutes les modalités d'une variable donnée.

- envoyer dans MS Office® le contenu d'un tableau de bord.



On peut ensuite imprimer, enregistrer au format HTML ou directement dans MS Office®.

Analyse en ligne

A partir d'une enquête menée sur Internet, vous avez la possibilité de suivre vos résultats en ligne. Pour cela, il vous faut tout d'abord préparer dans les **Tableaux de bord** les résultats que vous voulez obtenir : tableaux à plat, de groupes, croisés, graphiques, cartes etc.

Dans les tableaux de bord, choisir le menu **Fichier** puis **Publier sur Internet**, suivre les différentes étapes. A l'étape **Publication**, choisir :

- **réinitialiser les fichiers sur le serveur**, si vous souhaitez remettre à zéro les réponses sur le serveur.

- **publier les réponses locales** si vous souhaitez mettre en ligne des réponses déjà saisies sur votre ordinateur.

En plus des liens vers le questionnaire et les tables, vous obtiendrez un lien supplémentaire de type : http://www.sphinxonline.net/nomducompte/nomenquete/resultats.htm. C'est ce lien qui vous permettra de suivre vos résultats en ligne.

Toutes les fonctions, menus et dialogues

10. Organisation du logiciel

Les 3 stades

Au démarrage, indiquez ce que vous voulez faire :

- Faire une enquête par questionnaire,
- Analyser des données existantes,
- Etudier des textes



Le logiciel est construit sur les trois stades d'une enquête : définition du questionnaire, collecte des réponses et traitements.

Elaboration du questionnaire



Après la définition de l'identité de la nouvelle enquête, le panneau central affiche les principales fonctions : rédiger le questionnaire, l'imprimer et le mettre en page, créer des formulaires multimédia et les publier sur Internet (**Formulaires multimédia**).

Les 2 menus déroulants du haut de l'écran (**Rédiger** et **Organiser**) listent en détail toutes les fonctions disponibles à ce stade.

Le bouton **Tables** offre un raccourci vers la liste des variables et vers les données.

Collecte des réponses



Ce stade est pertinent pour

- la saisie manuelle des réponses en **Saisie rapide** ou par **Module opérateur**.
- La qualification et la gestion des données

Les menus déroulants **Saisir** et **Gérer** listent en détail toutes les fonctions disponibles.

Traitements et analyses



Le dernier stade offre les fonctions d'un puissant outil statistique et d'analyse de données sur lequel on arrive directement dans le cas de travail sur des données existantes ou des textes.

Toutes les fonctions disponibles sont activées en mode classique par les articles des menus **Dépouiller**, **Recoder**, **Analyser** et **Approfondir**. En cliquant sur **Tableaux de Bord**, on les retrouve dans l'environnement Multimédia et Datamining.

Les 2 environnements de travail

A chaque stade, vous pouvez travailler selon :
- un mode classique : vous choisissez par menu une fonction que le logiciel exécute en vous guidant.
- un mode multimédia : vous manipulez des objets que vous disposez et dont vous fixez les propriétés.

Le premier mode est d'un abord plus simple ou moins abstrait mais le second est beaucoup plus rapide, et bien plus efficace.

A vous de choisir en fonction de votre style, de vos aptitudes et de ce que vous avez à faire.

Certaines tâches ne peuvent être effectuées que dans l'environnement classique (créer un questionnaire, ajouter une question, gérer les données…), d'autres passent obligatoirement par l'environnement multimédia (faire un formulaire Internet, concevoir un tableau de bord…).

L'accès à l'environnement multimédia se fait par les boutons suivants :
- **Formulaires multimédia** au stade du questionnaire ;
- **Module opérateur** au stade de la collecte des réponses ;
- **Tableaux de bord** au stade des traitements.

A chaque stade une expertise

A chaque stade, vous disposez d'une expertise de votre travail :
- **Evaluation du questionnaire**
- **Evaluation du jeu de données**
- **Conseil pour l'analyse**

Elle est fondée sur un examen automatisé de votre travail et vous fait bénéficier de la base de connaissances de Sphinx.

L'atelier lexical

Un dernier panneau, disponible uniquement dans la version Lexica, conduit à l'environnement d'analyse lexicale et d'analyse de données textuelles. Depuis le panneau de traitement, on y accède par **Etudier les textes**.

L'environnement classique

L'environnement classique fonctionne par menus, fenêtres et boîtes de dialogue.

Les menus

Les menus disposés en haut de l'écran comportent une partie commune :

- Fichier : donne accès aux fonctions d'accès et de sauvegarde des fichiers (**Ouvrir, Enregistrer**, **Faire une sauvegarde**) et de sortie (**Imprimer**, **Inclure dans le rapport**…) ;

- Edition : **Copier**, **Coller**, paramétrages (**Polices**, **Langues**, **Couleurs**…) ;

- Stade : donne accès aux différents stades.

Les menus spécifiques apparaissent sur la droite selon le stade sélectionné.

Fenêtres et boîtes de dialogues

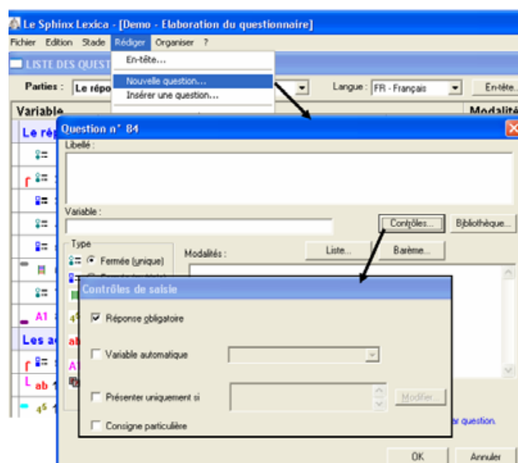
La sélection d'un article de menu ouvre une fenêtre dans laquelle s'élabore la tâche à

effectuer (questionnaire, saisies, tableaux…). L'action est lancée en cliquant sur des boutons ou en fixant des choix dans des boîtes de dialogue.

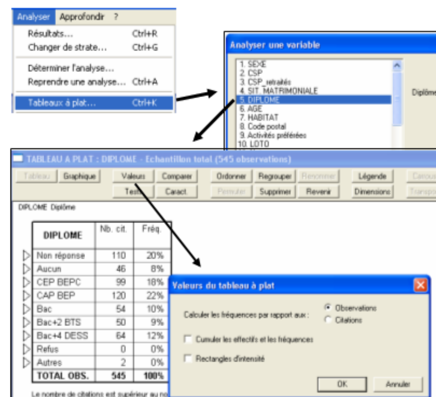
Par exemple au stade questionnaire :

- Rédiger, **Nouvelle question** ouvre la fenêtre de définition des questions.

- Le bouton **Contrôle** ouvre le dialogue de contrôle de saisie dans lequel on coche la case qui rend la réponse obligatoire.



Au stade des traitements, **Analyses**, **Tableaux à plat** ouvre une liste déroulante dans laquelle on choisit la variable à analyser. Le tableau apparaît dans la fenêtre **Tableaux à plat**. Le bouton **Valeurs** ouvre le dialogue de paramétrage des fréquences.



Enregistrement et sauvegarde

Dans l'environnement classique, Sphinx sauvegarde automatiquement la structure du questionnaire en cours d'élaboration. Les réponses saisies et les modifications de la table des données consécutives aux opérations de gestion des données (transformation de variables, recodification, calcul de nouvelles variables) sont également automatiquement enregistrées.

Les résultats des analyses sont eux « volatiles » et doivent être imprimés ou transférés dans un traitement de textes (**Fichier**, **Inclure dans le rapport**) pour être conservés.

Fichier, **Faire une sauvegarde**, **Exporter** permet de réaliser des copies d'archives ou de renommer les fichiers.

Les fichiers Sphinx

Une application du logiciel crée 4 fichiers portant le nom de l'enquête et distingués par leur extension :

- Le questionnaire : nom.que
- Les réponses fermées : nom.rep
- Les réponses ouvertes textes : nom.ouv
- Les formulaires et tableaux de bord : nom.hyp

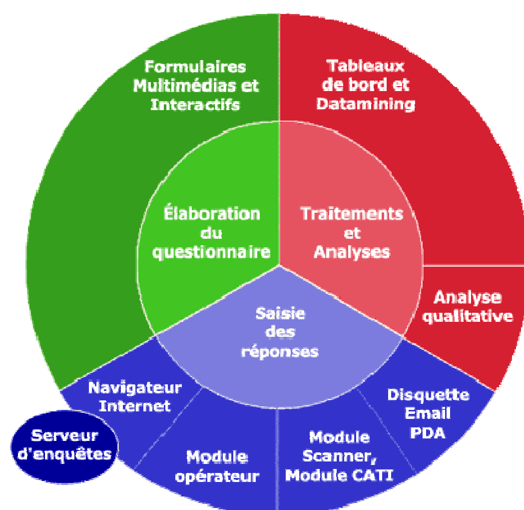
Relation avec l'environnement multimédia

Le passage par l'environnement classique est obligatoire pour la création du questionnaire ou l'ajout de questions supplémentaires.

Pour publier un questionnaire sur Internet, préparer un questionnaire à scanner, faire du CATI ou personnaliser un formulaire, le passage par l'environnement multimédia est obligatoire.

Tout le reste et notamment le travail de dépouillement ou d'analyse peut être fait indifféremment dans l'un ou l'autre des environnements.

Le schéma ci-dessous montre les 3 stades avec au centre les parties correspondant à l'environnement classique et autour la couche multimédia. Le passage d'une zone à l'autre se fait par le panneau central auquel on revient avec la touche **Esc**. ou en fermant l'environnement multimédia.



L'environnement multimédia

L'environnement multimédia fonctionne en mode objet, ce qui lui confère sa puissance comme :

- outil de présentation : les objets (questions, variables, légendes, illustrations) sont disposés dans des vues. Leurs propriétés modulables à souhait fixent les contenus et les aspects. On peut ainsi construire des formulaires adaptés à tout type de média.
- solution de productivité pour l'analyse de données : une fois fixés, les vues et tous les objets qui les composent, peuvent être appliqués à tout nouvel ensemble de données en reconstituant instantanément les analyses qu'ils contiennent.

A partir du panneau central, on accède à ce mode de fonctionnement par les boutons :

- **Formulaires multimédia** au stade Questionnaire ;
- **Tableaux de Bord** au stade Traitement.
-
- Dans l'environnement multimédia, on passe des formulaires aux tableaux de bord par le menu **Stade** : **Définir les formulaires** ou **Explorer et analyser**

L'espace de travail multimédia

L'organisation est la même pour les formulaires ou les tableaux de bord.

La fenêtre des vues

Une vue est composée de pages dans lesquelles sont disposées les questions pour les formulaires ou les tableaux et graphiques pour les tableaux de bord.

La liste des variables

Cette liste permet de sélectionner les variables à placer dans la vue (double clic ou clic glissant). Les sélections peuvent également se faire par menu (**Insertion** pour les formulaires ou **Analyses** pour les **Tableaux de Bord**) ou par les raccourcis de la barre d'outils.

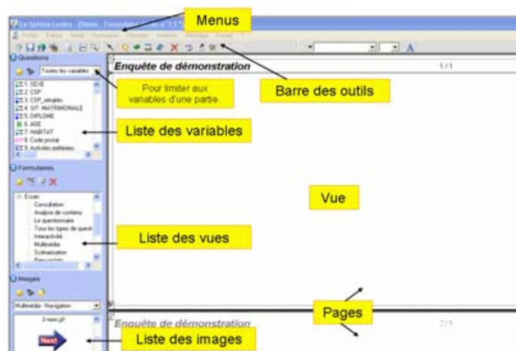
La liste des vues

Selon les média utilisés pour collecter les données, on peut préparer différentes mises en forme réalisées dans différentes vues. La liste des vues permet de les choisir et de les « monter » dans la fenêtre des vues. De même, on pourra mener les analyses sur plusieurs tableaux de bord.

La liste des strates

Cette liste n'est disponible que pour les tableaux de bord. Elle donne accès aux profils selon lesquels on peut voir les résultats de la vue active.

La liste des images



Une banque d'images est proposée avec le logiciel pour l'illustration des graphiques, notamment pour la construction de barres de navigation personnalisées.

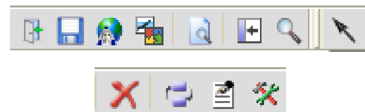
Les menus

Dans le haut de la fenêtre sont disposés les menus. Ils sont spécifiques à chaque stade et portent principalement sur :

- l'insertion d'objets (**Insertion**, **Analyses**) ;
- la disposition des objets (**Affichage**) ;
- l'aspect et le style des objets (**Format**) ;
- l'accès aux données brutes (**Données**).

Le bandeau des outils

- Commun aux 2 stades :



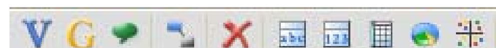
Fermer, Sauver, Publier sur Internet, Envoyer dans Office, Aperçus, Zoom, Sélectionner, Supprimer, Placer, Propriétés

- Formulaires :



Insérer une question, une légende, une image, une vidéo

- Tableaux de bord :

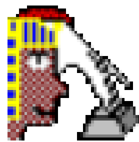


Insérer une variable, un groupe, une légende, une relation, supprimer.

Afficher, faire disparaître le titre, la description, le tableau le graphique, le mapping.

Le module opérateur

Le module opérateur est une version du logiciel limitée aux fonctions de saisie / modification des réponses et de visualisation des résultats. Il peut être configuré selon les usages souhaités et installé sur le poste de la personne à laquelle il est destiné.



Il est lancé comme logiciel indépendant et ouvre le panneau de commandes suivant.

Enquêtes décentralisées

Prenons l'exemple du service central d'une grande entreprise qui charge les agences de son ressort de lancer périodiquement des enquêtes de satisfaction. Avec le module opérateur, les agences peuvent :

- saisir les réponses papier ou importer celles venant d'Internet. L'opérateur de saisie choisit le masque qui lui convient ou opte pour une saisie scanner.
- consulter les réponses à des fins de vérification ou d'édition de fiches.
- voir les résultats dans les tableaux de bord préparés à cet effet par le centre.

Voici le panneau de commandes du module opérateur.



On aurait pu également décider de ne pas autoriser le retour sur les réponses et ne faire apparaître que le bouton **Consulter/Modifier**.

Le module opérateur simplifie la tâche des agences tout en leur laissant l'initiative du travail. Le centre est garanti que toutes les agences utiliseront le même modèle d'enquête et qu'elles prendront connaissance des résultats conformément à sa vision contenue dans les tableaux de bord.

Modules de saisie

Le module opérateur est fréquemment utilisé pour équiper des postes de saisie. Le panneau de commande ne comporte alors que le bouton **Saisir**.

(*) L'ajout du bouton **Importer** permet la consolidation des saisies opérées sur les différents postes.

Opérateur CATI (*)

Le système CATI intègre un ensemble de postes téléacteurs partageant une base d'appels commune et un fichier de campagne. Il met en œuvre un ensemble de modules opérateurs pilotés par des fonctions de supervision.

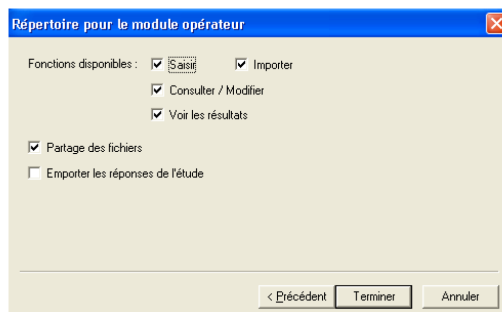
Affectation d'une enquête au module opérateur (*)

Le module opérateur doit être installé sur le poste de la personne à laquelle on le destine.

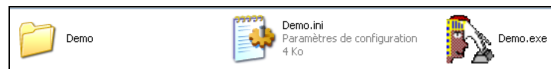
Une fois l'enquête mise au point (formulaires et tableaux de bord), on la prépare pour le module opérateur :

Environnement multimédia / Fichier / Emporter l'enquête / Répertoire pour le module opérateur

En fixant les fonctions à autoriser :



Les 3 éléments ci-dessous sont créés pour être installés sur le poste de travail du destinataire.



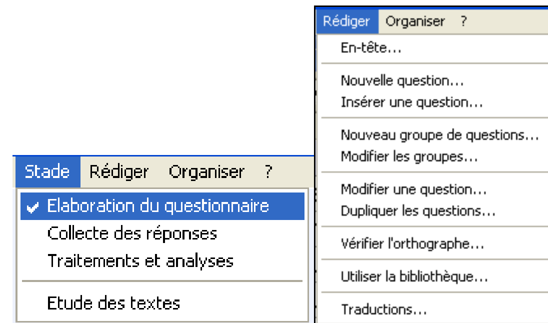
Le dossier contient les fichiers de l'enquête. L'icône est celle du lanceur. Il démarre automatiquement le module opérateur sur l'enquête contenue dans le dossier.

11. L'environnement classique

Stade Questionnaire

Menu Rédiger

Pour définir les éléments du questionnaire : **En-tête**, **Nouvelle question**.



Les questions ou groupes de questions s'ajoutent dans la fenêtre **Liste des questions**.

Insérer une question insère la question avant l'élément sélectionné dans la liste.

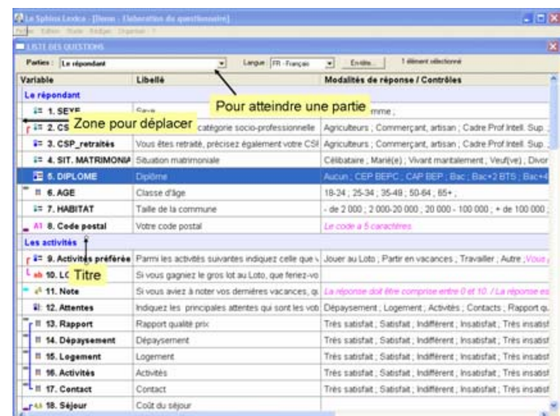
L'article **Nouveau groupe de questions** accélère la construction des groupes.

Vérifier l'orthographe pour corriger le questionnaire en cours de construction.

Utiliser la bibliothèque propose des exemples de questions.

Liste des questions

La rédaction s'élabore dans la fenêtre de définition des questions.



La largeur des colonnes est modifiable en faisant glisser le trait de séparation.

Pour modifier un titre ou la question : double clic sur la ligne correspondante.

Pour déplacer les questions ou titres : cliquer dans la zone de droite, et faire glisser la ligne sélectionnée.

Définition des questions

La rédaction des questions s'élabore dans la fenêtre de définition des questions.

Une zone apparaît pour la définition des modalités des questions fermées. Le bouton **Liste** permet d'en modifier l'ordre.

Contrôles pour spécifier des conditions supplémentaires (réponse obligatoire, condition de présentation, borne pour les numériques…).

Barème : pour associer une valeur aux modalités en les entrant en face de chaque modalité ou en les énumérant (**Edition auto**…).

Permet notamment de donner une valeur numérique aux échelons des variables échelles.

Contrôles : pour

- rendre la question obligatoire ;
- lui associer une question « Autre préciser » ;
- conditionner son apparition par un profil ;
- lui associer une valeur automatique : une clé (pour revenir et modifier des réponses sur Internet), une date, une valeur identique à la saisie précédente, le résultat d'un calcul, un compteur…

Utiliser la bibliothèque

La bibliothèque contient des exemples de questions répertoriées par thème.

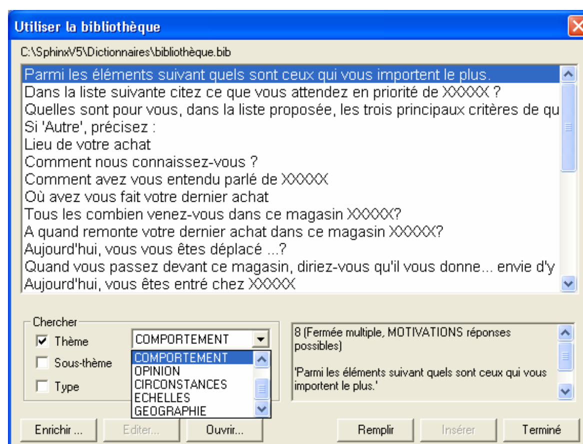
Remplir : pour insérer dans le questionnaire en cours de construction.

Les questions sont contenues dans une base Sphinx que l'on peut modifier.

Enrichir : pour ajouter une question de l'enquête en cours à la bibliothèque.

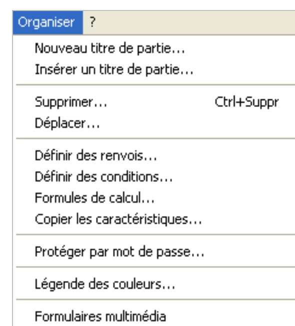
Editer : pour ouvrir le fichier de la bibliothèque. On quitte l'enquête en cours.

Ouvrir : pour utiliser comme bibliothèque une enquête existante.



Menu Organiser

Pour insérer des parties, définir des renvois et restrictions…



Renvois et restrictions

Renvois

pour sauter un ensemble de questions

1/ **Après la question** : choisir la question qui lance le renvoi.

2/ **Définir le profil** : définir les conditions du renvoi.

3/ Fixer la destination (**Aller** à) : choisir la question de destination.

Conditions de présentation :

Pour contrôler l'affichage d'une question en fonction des réponses précédentes : choisir la question qui définit la condition et les modalités ou plage du profil. Répéter en créant une condition par ligne.

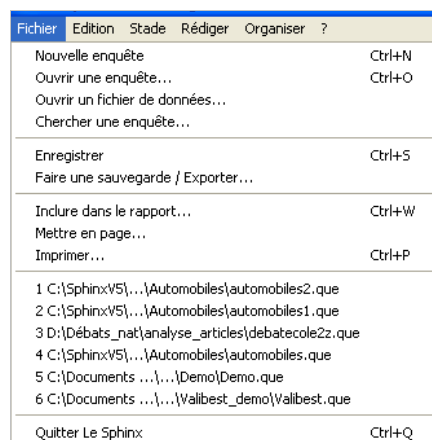
Formules de calculs (*)

Formules de calculs… :

Pour revoir et/ou modifier l'ensemble des formules de calcul des variables calculées.

Menu Fichier

Créer une nouvelle enquête, enregistrer l'enquête en cours, ouvrir et retrouver des enquêtes existantes.



Mettre en page, imprimer et transférer le questionnaire dans un traitement de textes.

La liste des dernières enquêtes utilisées apparaît en fin de menu.

Mettre en page

Pour imprimer directement un questionnaire bien présenté ou le mettre dans un traitement de textes.

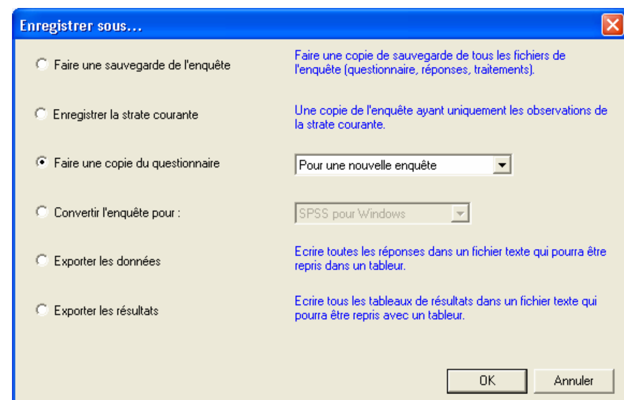
Mettre en page :

Pour choisir les options de présentation : Disposition, En-tête, Numérotation, Zones, Polices, Réponses

Inclure dans le rapport :

Pour envoyer la question dans un traitement de textes.

Enregistrer sous…



Pour :

Faire une sauvegarde du questionnaire avec (**Faire une sauvegarde de l'enquête**) ou sans les données (**Faire une copie du questionnaire**).

Enregistrer un sous ensemble des données (strate) : **Enregistrer la strate courante**

Convertir l'enquête Sphinx vers un autre format : Spss, Triple-s…

Stade Collecte des réponses

Pour saisir ou importer les réponses et gérer les fichiers de données.

Menu Saisir

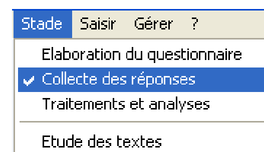
Saisie directe : question par question

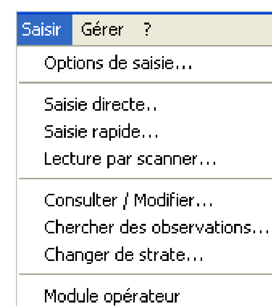
Saisie rapide : pour une saisie au kilomètre optimisée (options de saisie)

Module opérateur : pour saisir dans l'environnement multimédia

Lecture par Scanner : lance le module de lecture par scanner pour lire automatiquement le formulaire préparé pour le scanner.

Consulter/modifier : pour parcourir la base de données et faire des modifications (mode tableur, fiche ou direct).





Saisie rapide

Pour une saisie au kilomètre : la solution la plus efficace pour entrer une pile de questionnaires.

Ajouter une modalité : à manier avec précaution car ceci affecte la structure du questionnaire.

Options : Pour régler le fonctionnement de la saisie rapide et accélérer la saisie des réponses ouvertes.

Consulter/modifier

Pour parcourir les réponses et éventuellement les modifier.

Mode direct : question par question (icône en bas à gauche).

Mode rapide : analogue à la saisie rapide.

Tableur : pour voir, modifier ou saisir en mode tableur de la manière suivante.

Strates | Variables | Observations | Fiches | Liste de valeurs

Toutes les variables | 1;3;26;25 | Noms des modalités

N.	1. POSSESSION	3. MARQUE	26. SEXE	25. VOITURE IDEALE
1	Oui	Citroën	Femme	Elle est rouge, petite et tote ronde.
2	Oui	Citroën	Homme	Une limousine avec un bar
3	Oui	Autre	Homme	Une bonne voiture de ville
4	Oui	Renault	Homme	Voiture confortable
5	Oui	Peugeot	Femme	Mercédes
6	Oui	Renault	Femme	une limousine spacieuse et raffinée
7	Oui	Renault	Homm	
8	Oui	Peugeot	- / Homme	vec des chromes éclatants , super-rapide
9	Oui	Ford	Femme	

Afficher les données : toutes les variables, une seule (choisie dans la liste) ou quelques-unes, définies en tapant la liste de leurs numéros séparés par ; (ou - pour toutes les variables d'une plage : 7-11 par ex.)

- Modifier les valeurs par clic sur la cellule ;

- Trier selon l'ordre croissant ou décroissant des valeurs (clic sur le titre de colonne) ou selon l'ordre d'enregistrement (colonne N°) ;

- Restreindre à une strate ou un profil.

Module opérateur (*)

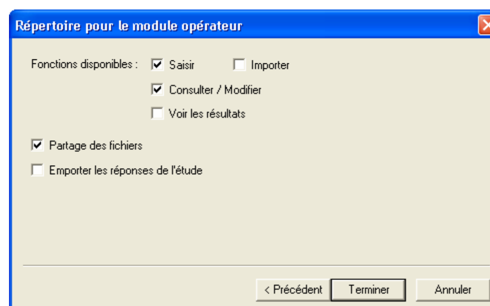
Pour créer un écran de saisie personnalisé et sécurisé.

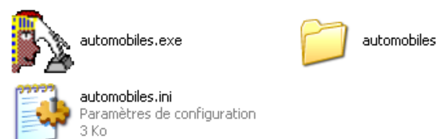
1/ Définir le formulaire dans la partie multimédia.

2/ **Emporter l'enquête** (menu **Fichier**) en précisant les droits (saisir, modifier…).

3/ Installer les fichiers sur les machines où doit se faire la saisie.

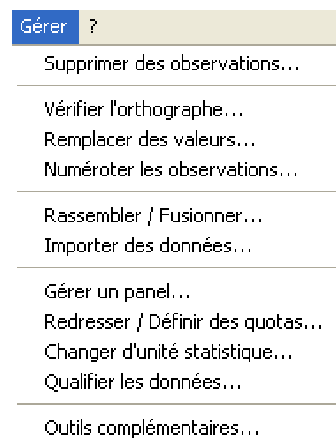
4/ Le lanceur donne directement accès à l'écran de saisie.





Menu Gérer

Gérer le fichier des réponses



Supprimer des observations : pour nettoyer les données (supprimer les observations à non réponses : Les observations telles que...),

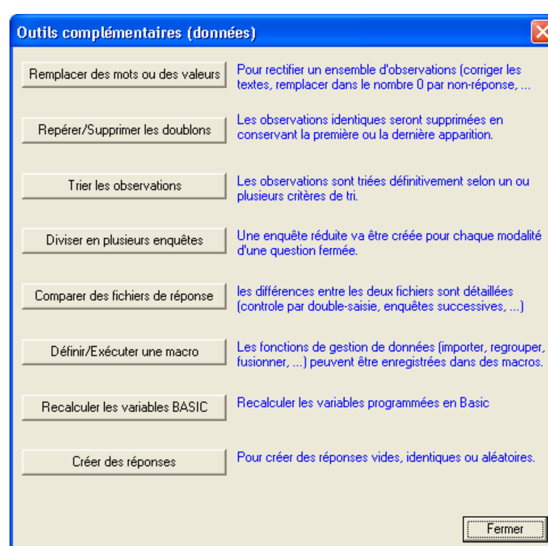
(*) Remplacer des valeurs et/ou expurger des réponses textes en supprimant les mots d'une liste donnée. Sur toutes les données ou sur une strate seulement.

(*) Vérifier l'orthographe (question texte) **/, Numéroter les observations** en créant une nouvelle variable numérique contenant le numéro de l'observation.

(*) Augmenter la base de données : **Rassembler/Fusionner** les données provenant de différentes saisies ou enquêtes, **Importer des données** saisies via Internet ou provenant de fichiers externes.

(*) Transformer : **Gérer un panel, Redresser/Définir des quotas, Changer d'unité statistique, Qualifier les données**. (voir détail ci-dessous).

(*) Outils complémentaires : un ensemble de fonctions pour intervenir rapidement sur les données.



(*) Rassembler, Importer

(*) Rassembler / Fusionner s'applique à des données collectées avec Sphinx alors qu'importer consiste à transformer des données en provenance d'une autre application en fichiers Sphinx.

(*) Rassembler / Fusionner : se place dans une enquête Sphinx à laquelle on vient ajouter les données d'autres enquêtes.

(*) Importer des données : on part d'un questionnaire défini dans Sphinx dans lequel on importe des données.

L'assistant d'importation guide l'opération :

1/ indiquer les données à importer,

2/ fixer les séparateurs,

3/ affecter les colonnes du tableau à une variable,

4/ contrôler la qualité des données,

5/ fixer les modalités de création de la base Sphinx en précisant si les données importées remplacent ou viennent s'ajouter aux données existantes.

Pour récupérer dans Sphinx des données externes :

(*) Ouvrir un fichier de données (Fichier) : Ne part pas d'un questionnaire existant ; c'est à l'étape 3 qu'on définit le questionnaire en précisant le nom et le type des variables.

Importer les fichiers de réponses sur Internet : entrer le nom du compte SphinxOnline et celui de l'étude.

Rassembler / Fusionner

Rajouter des observations d'une enquête identique

Une partie de la saisie a été effectuée sur un autre poste. La fonction consiste à ajouter ces réponses à la fin de l'enquête courante. On indiquera le fichier des réponses (*.rep) à rassembler.

Rassembler les variables communes de plusieurs enquêtes

Différentes enquêtes similaires comportent des variables communes. Une nouvelle enquête reprendra ces variables pour l'ensemble des observations de toutes les enquêtes.

Fusionner deux enquêtes

Ajouter à l'enquête en cours des variables d'une autre enquête que l'on indiquera. Une variable 'clef' sera demandée pour assurer l'identification d'une même observation dans les deux enquêtes.

OK | Annuler

Méthode d'importation

Importer des données — Un fichier (txt, xls, mdb, ...) ou une source de données ODBC va être utilisé pour créer de nouvelles observations ou modifier/compléter les observations existantes.

Ouvrir un fichier de données — Un fichier (txt, xls, mdb, ...) ou une source de données ODBC va être analysé pour en déterminer la structure. Une nouvelle enquête sera créée.

Importer des fichiers de réponses sur Internet — Télécharger les réponses saisies sur SphinxOnline ou un autre serveur d'enquêtes.

OK | Annuler

Importer - Etape 2/5

Ignorer les 0 premières lignes

La première ligne utile contient le nom des variables

Séparateurs entre champs : Tabulation

Séparateurs entre observations : Retour chariot

Caractère délimiteur : "

Caractères pour les non-réponses :

51 colonnes, 202 lignes

	1. POSSESSION	2. APPARTENANCE	3. MARQUE	4. MARQUE_AUTRE	5. TYPE AUTO
2	1	4	3		2
3	1	5	5		2
4	1	6	12	Ferrari	1
5	1	7	1		1
6	1	8	2		2
7	1	7	1		1
8	1	7	1		1
9	1	8	2		2
10	1	9	6		1
11	1	10	9		2
12	1	7	1		2
13	1	11	4		1
14	1	12	10		2

< Précédent | Suivant > | Annuler

Redresser, échantillonner, qualifier

Redresser, définir des quotas

Choisir la ou les variables pour lesquelles on veut respecter des quotas.

Choisir un type de redressement : extraction, pondération, substitution.

Aide à l'échantillonnage :

Pour connaître la précision des estimations en fonction de la taille de l'échantillon courant ou déterminer la taille de l'échantillon nécessaire pour réaliser un objectif de précision.

Définir des quotas

Quota 1
Quota 2
Quota 3
26. SEXE
0.50 : Non réponse
48.76 : Homme
50.75 : Femme
48.76 Homme
Méthode de redressement
Extraire un échantillon
Une nouvelle variable 'REDRESSEMENT' sera utilisée pour identifier un sous-ensemble des observations respectant les quotas.
Pondérer les observations
Une nouvelle variable 'REDRESSEMENT' sera utilisée pour pondérer les effectifs dans les tableaux à plat et les tableaux croisés.
Substituer des observations
Une nouvelle enquête sera créée en subsituant des observations pour respecter les quotas.
Fixer l'effectif redressé
201
Résultats...
Aide à l'échantillonnage...
OK
Annuler

(*) Qualifier les données :

- Remplacer les non-réponses par la valeur moyenne ou le mode, la valeur précédente ou le plus proche voisin.
- Repérer les observations singulières par rapport à leur(s) valeur(s) sur une ou plusieurs variables ou par leur nombre de non-réponses.

(*) Changer d'unité statistique :

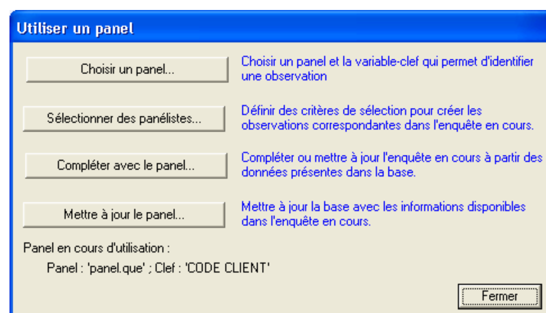
Créer une nouvelle enquête en éclatant les données selon les éléments répétés de la base de départ (parties, nombres..) ou en regroupant les observations concernant le même individu.

Qualifier une enquête : identifier les observations singulières

L'observation a une valeur singulière pour
Une des variables sélectionnées
Toutes les variables sélectionnées
1. POSSESSION
2. APPARTENANCE
3. MARQUE
4. MARQUE_AUTRE
5. TYPE AUTO
6. ANCIENNETE
0
1 élément sélectionné
Une valeur est dite singulière si elle est beaucoup trop éloignée de la moyenne (>3.56 s) ou bien dans une catégorie de trop petit effectif (<1%).
L'observation n'est pas suffisamment documentée
Une observation est considérée comme insuffisamment documentée si 50% ou plus des questions sont à non-réponse.
Une variable va être créée permettant de construire une strate sur les "bonnes observations" ou bien de supprimer définitivement les "mauvaises observations".
OK
Annuler

Gérer un panel (*)

Pour associer à une enquête (la campagne), les données provenant d'un fichier existant (le panel).

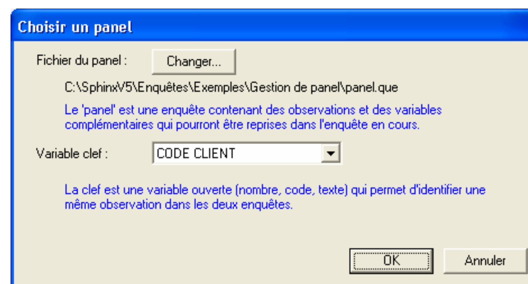


Choisir un panel :

Permet d'associer à la campagne le fichier panel.

Sélectionner des panélistes :

Ajouter dans l'enquête les données provenant du panel pour les utiliser au moment de la collecte (CATI, Internet).

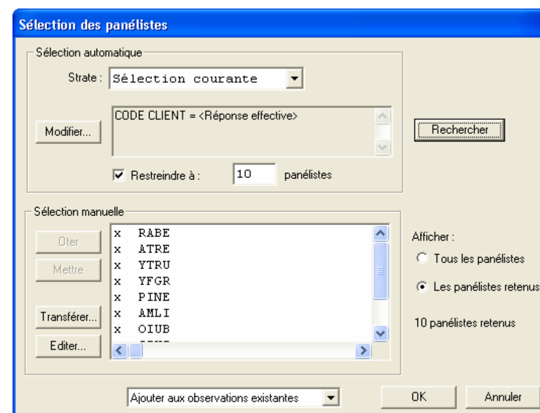


Compléter avec le panel :

Pour ajouter aux données collectées les informations provenant du panel.

Mettre à jour le panel :

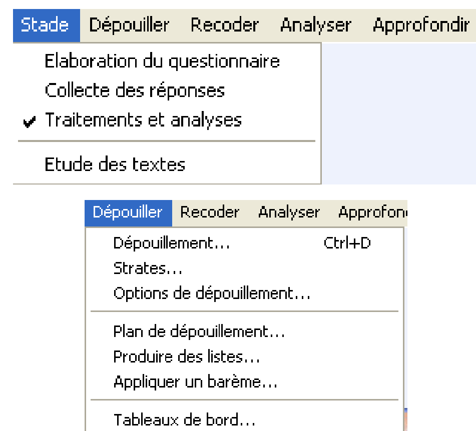
Pour corriger le panel avec les informations collectées pendant la campagne.



Stade Traitements

Traitements et analyses

Dans l'environnement classique, les traitements sont exécutés à la demande sans possibilité de les conserver ni de les mettre en page autrement qu'en les exportant vers un traitement de textes.



Dépouiller : lancer d'un seul coup tous les traitements et produire un rapport.

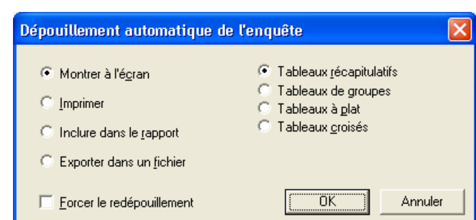
Recoder : mettre en classes les numériques, interpréter les réponses ouvertes, les dates et les codes, transformer les variables fermées.

Analyser : exécuter les tris à plat et croisés et adapter la présentation des résultats.

Approfondir : mettre en œuvre les méthodes avancées d'analyse multivariée.

Dépouiller

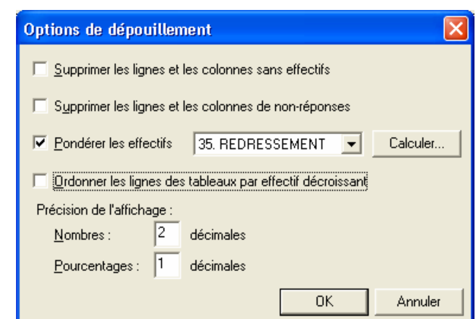
Modes de dépouillement



Dépouillement : pour lancer tous les traitements sélectionnés et les faire défiler à l'écran, les imprimer, les inclure dans un rapport ou les exporter.

Inclure dans un rapport : pour envoyer directement le rapport dans un document Word.

Strates : pour choisir la strate sur laquelle effectuer le dépouillement.



Options de dépouillement : pour fixer les modalités du dépouillement. La présentation des tableaux correspond aux options courantes du menu d'analyse.

Pondérer les effectifs nécessite de fixer les critères de redressement et le calcul (**Calculer...**) de la variable contenant les poids (REDRESSEMENT). (Voir aussi **Redresser/Définir des quotas** au stade **Collecte des réponses**).

Plan de dépouillement

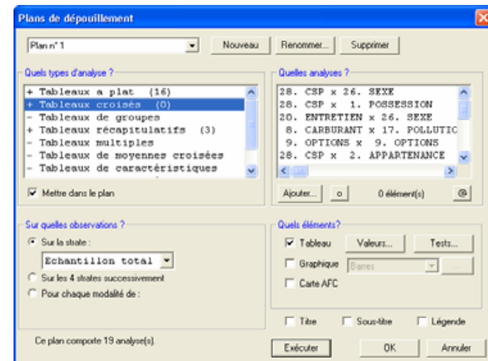
Pour définir un ensemble de traitements qui pourra être exécuté à nouveau après avoir été enregistré.

Fixer :

- les types de traitements à inclure (Tableaux à plat, Tableaux croisés, Tableaux de groupes…),

- la présentation choisie (Tableau, Graphique…),

- les données sur lesquelles lancer le plan (échantillon total, strate(s), catégories définies par les modalités d'une variable).

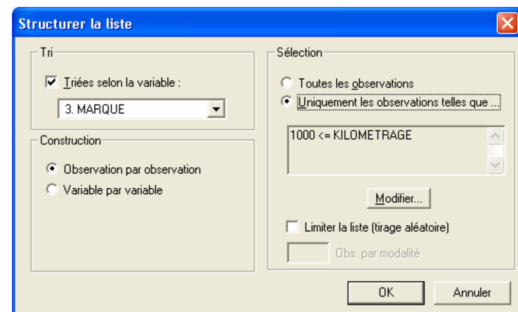


Produire des listes

1/ Choisir les variables à inclure dans la liste.

2/ Définir la présentation (ordre des variables, format fiche ou liste).

3/ Définir la structuration éventuelle : tri de la liste et/ou restriction à un profil.

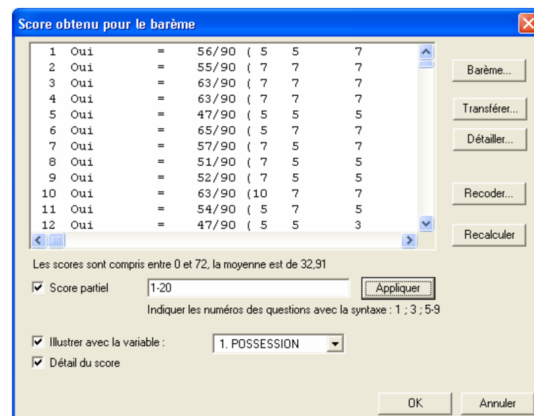


Appliquer un barème

Pour créer une nouvelle variable contenant la valeur du barème éventuellement réduite à un sous-ensemble de variables (score partiel).

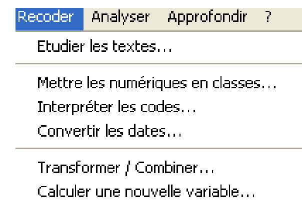
Détailler : pour vérifier le calcul du barème.

Recoder : pour créer une variable contenant le résultat du calcul.



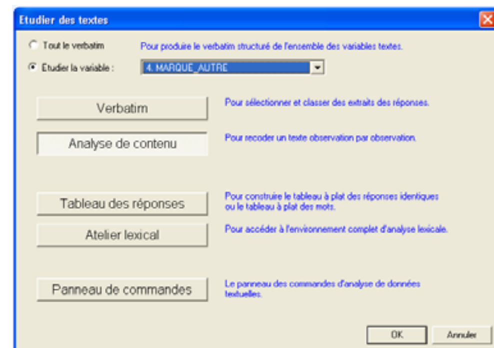
Menu Recoder

Transformer et enrichir les données existantes en créant de nouvelles variables.



Etudier les textes

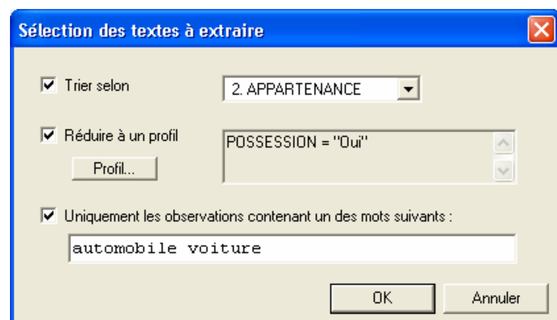
- restituer tout ou partie du texte des réponses (**Verbatim)**,
- lire, interpréter et coder les réponses en utilisant une grille thématique (**Analyse de contenu**),
- Etudier statistiquement la composition du texte des réponses (**Tableau des réponses, (*) Atelier lexical**).



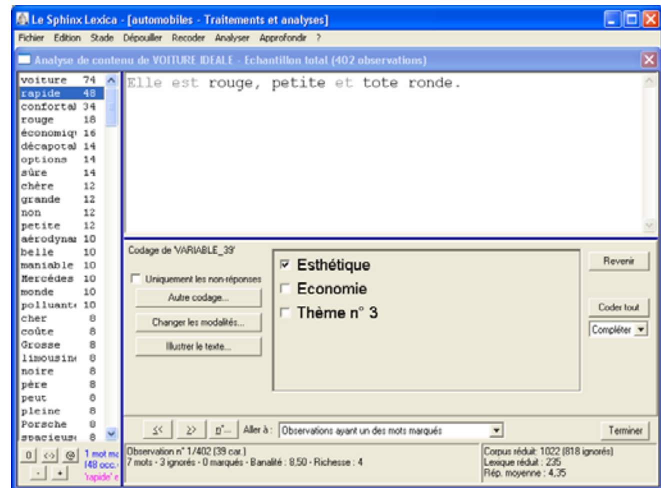
Verbatim

Pour lister la totalité ou une sélection d'extraits des réponses en :

- les triant par catégorie d'une variable de contexte (**Trier selon**),
- en se concentrant sur les réponses répondant à un profil particulier (**Réduire à un profil**) ou contenant certains mots (**uniquement les observations contenant un des mots suivants**…).



Analyse de contenu



Pour créer une ou plusieurs nouvelles variables dans lesquelles coder, après lecture, les idées contenues dans les réponses.

1/ créer la nouvelle variable destinée à coder les idées présentes dans le texte,

2/ définir les éléments de la grille de lecture (**Changer les modalités**),

3/ parcourir les réponses et cocher en fonction du contenu,

4/ éventuellement créer une autre variable pour coder selon un autre point de vue (**Autre codage**),

5/ passer en revue l'ensemble des réponses en utilisant les flèches pour passer d'une réponse à l'autre,

6/ sélectionner des mots dans le lexique pour ne voir que les réponses contenant ces mots, (**Coder tout** pour cocher automatiquement les réponses contenant les mots sélectionnés).

Attention : cet environnement est utile pour mettre au point les grilles de codification. Pour le travail de lecture et de codification proprement dit, il est plus intéressant de préparer un formulaire multimédia pour un meilleur confort de travail dans le module opérateur.

Tableau des réponses

Quand les réponses sont brèves (autres, préciser notamment), le tableau des réponses donne rapidement une idée du contenu. L'atelier lexical permet d'aller beaucoup plus loin pour les textes élaborés et volumineux.

Mots : pour voir la liste des mots et leur fréquence (**Réponses** pour les réponses entières).

Recoder : pour créer une variable réduite aux mots sélectionnés.

Croiser : pour croiser avec une autre variable.

Le Sphinx Lexica - [automobiles - Traitements et analyses]

Fichier Edition Stade Dépouiller Recoder Analyser Approfondir ?

TABLEAU DES REPONSES - Echantillon total (402 observations)

Croiser Valeurs Mots Transférer

Analyser Recoder Tests Caract. Ordonner... Supprimer Regrouper Revenir Autre

VOITURE IDEALE - Décrivez la voiture idéale selon vos désirs

Valeurs	Nb. cit.	Fréquence
voiture de sport	6	2,1%
confortable	4	1,4%
économique	4	1,4%
Seule la qualité compte à mes yeux. Une allemande solide et sûre	4	1,4%
tôle froissée	4	1,4%
/ grande/ rapide/ confortable rouge ou jaune	2	0,7%
8 grandes places	2	0,7%
aérodynamique	2	0,7%
aérodynamique, non polluante	2	0,7%
Amphibie mi-auto mi-bateau, avion aussi	2	0,7%
anti-PV	2	0,7%
avec des chromes éclatants , super-rapide	2	0,7%
avec un tableau de bord en bois laqué et un chauffeur en tenue	2	0,7%
avec une jolie fille dedans	2	0,7%

Nombre de valeurs différentes : 137
'voiture de sport' est le plus cité : 6 observations.
Il y a 116 non-réponses.

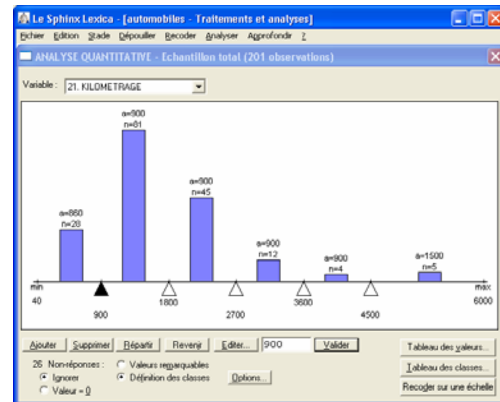
Mettre les numériques en classes

Pour analyser les variables numériques sous forme de tableau :

1/ choisir la variable

2/ fixer les classes d'intervalles en jouant sur le nombre et la position des seuils, ou en les entrant au clavier (**Editer**),

3/ voir le tableau (**Tableau des classes**) ou créer une nouvelle variable construite sur ces classes (**Recoder sur une échelle**).



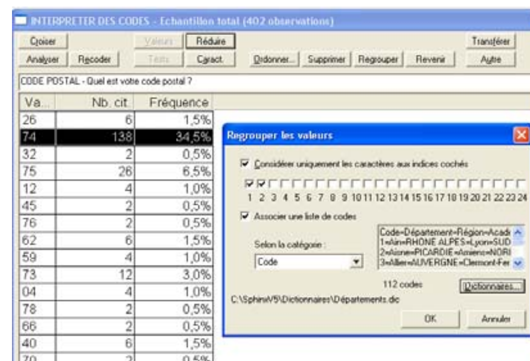
Interpréter les codes

Pour remplacer un code alphanumérique par les éléments auxquels il renvoie. :

1/ indiquer les positions significatives du code (les 2 premières pour le code postal),

2/ si cela n'a pas été fait au stade du questionnaire, associer le dictionnaire contenant les significations du code,

3/ actionner **Recoder** pour créer une variable contenant le code interprété.



Convertir les dates

Pour remplacer une date complète par l'année, le mois, le jour de la semaine, l'heure ronde… ou calculer une durée : indiquer une variable ou une valeur fixe pour les dates de début et de fin.

Recoder pour créer la nouvelle variable correspondant à la conversion.

Transformer une variable (*)

Pour créer de nouvelles variables à partir des variables de l'enquête :

-**Transformer une variable**

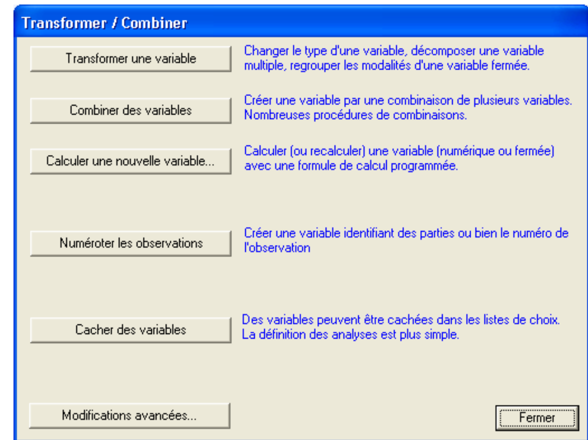
-Combiner des variables

-Calculer une nouvelle variable

-Numéroter les observations

-Cacher des variables

-Modifications avancées



Les calculs sont lancés et la nouvelle variable créée. A l'exception de **Calculer une nouvelle variable**, toutes les autres transformations devront être relancées pour intégrer de nouvelles données.

Transformer une variable

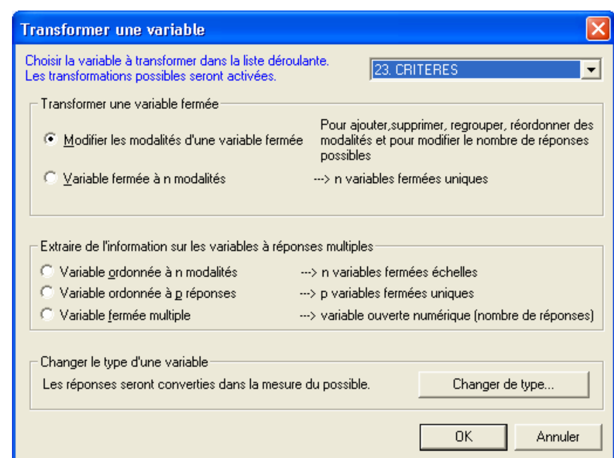
Par modification, regroupement des modalités d'une variable fermée ou par éclatement en n variables oui/non

Extraire de l'information sur les variables multiples

Pour exploiter le rang des modalités dans les variables fermées ordonnées ou calculer le nombre de réponses différentes des variables fermées multiples.

Changer de type :

Pour modifier le type d'une variable.



Combiner des variables (*)

Pour créer une nouvelle variable à partir de données en provenance de plusieurs autres variables.

Fusion :

- simple : la nouvelle variable a toutes les modalités des variables à fusionner.

- composée : la nouvelle variable a les modalités de la première variable. Les modalités des autres variables viennent s'y ajouter.

- croisée : les modalités de la nouvelle variable résultent du croisement des modalités des variables choisies.

Combinaisons : pour identifier les combinaisons les plus fréquentes dans les réponses à une variable fermée multiple.

Combiner des variables

SCORE / BAREME	Le 'score' obtenu pour un ensemble de questions muni d'un barème.
PROFILS	Une variable fermée dont les modalités sont choisies parmi les strates définies.
FUSION	Une variable fermée, combinant les modalités de deux variables.
MULTI_OUINON	Une variable fermée multiple, ayant pour modalités un ensemble de variables 'oui/non'.
VERBOSE	Une variable ouverte texte réunissant les réponses à plusieurs questions de différents types.
IDENTIFICATEUR	Une variable de type code, identifiant une observation par son numéro et/ou par les réponses à différentes variables.
DUREE	Une variable numérique, différence en jours (ou en heures) de deux dates.
MOYENNE	Une variable numérique, moyenne des valeurs effectives de plusieurs variables numériques ou échelles.
QUALITE	Une variable numérique représentant le nombre de questions effectivement renseignées.
TRANSPOSE	Un groupe de P variables à N modalités, transposées d'un groupe de N variables à P modalités.
COMBINAISONS	Une variable dont les modalités sont les meilleures combinaisons des modalités d'une variable fermée multiple.

OK Annuler

Outils complémentaires (*)

Outils complémentaires

Ces fonctions permettent de :

- nettoyer les fichiers de données (**Simplifier les questions textes**, **Enlever les mises en classes**, **Supprimer tous les attributs**),

- accélérer certaines tâches (**Dichotomiser les questions fermées** en une seule fois, **Editer le fichier des textes**),

- faciliter l'importation de données relatives à des questionnaires réalisés sans passer par Sphinx (**Importer des libellés**).

Outils complémentaires (traitements)

Simplifier les questions multiples	Supprime les colonnes inutilisées dans les questions multiples.
Importer des libelles	Les libellés sont décrits dans un fichier texte avec la syntaxe variable <tabulation> libellé.
Dichotomiser les questions fermées	Chaque modalité de question fermée donnera lieu à une nouvelle question 'Oui ; Non ;'.
Enlever les mises en classes	Les mises en classes sont toutes réinitialisées aux valeurs par défaut (6 classes de même amplitude).
Editer le fichier des textes	Pour faire des opérations de recherche/remplacement sur l'ensemble des textes.
Supprimer tous les attributs	L'enquête est réouverte en la simplifiant : ne sont conservés que les questions et les réponses.

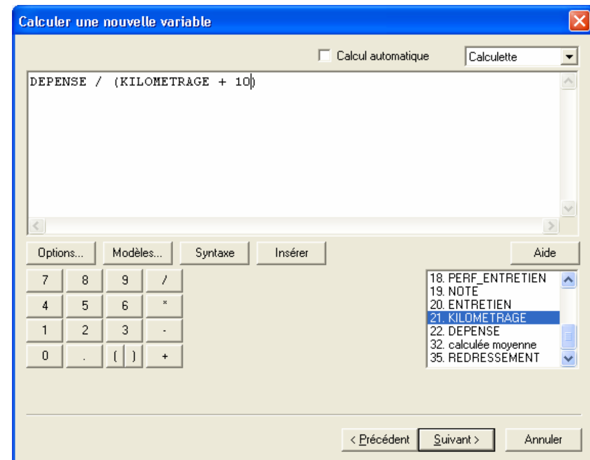
<< Précédent Fermer

Calculer une nouvelle variable (*)

Calculette

Pour écrire la formule de calcul d'une nouvelle variable.

En cochant **Calcul automatique**, le calcul sera actualisé à chaque nouvelle saisie.



Le bouton **Options** permet de traiter le cas des non-réponses de l'une des manières suivantes :

- en les ignorant : les observations dont l'un des éléments du calcul sont à non-réponse sont ignorées ;

- en les remplaçant par la valeur 0 ;

- en les remplaçant par la moyenne ou le mode.

Editeur assisté

Pour définir la nouvelle variable en utilisant les fonctions de SphinxBasic.

Les fonctions proposées sont réparties entre :

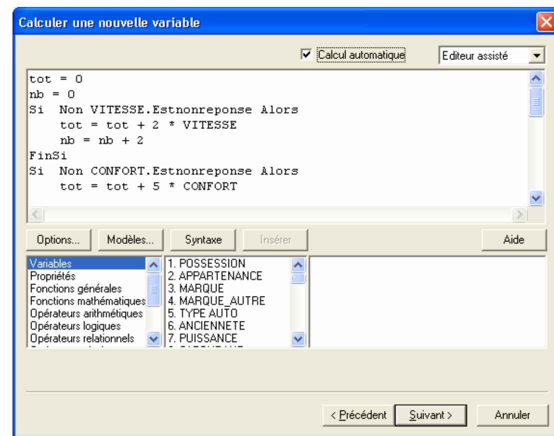
Propriétés : elles ne concernent qu'une seule variable (valeur, nombre de réponses effectives…).

Fonctions générales : elles s'appliquent à l'ensemble des données.

Opérateurs mathématiques ou logiques : pour définir les calculs.

Opérateurs chaînes et dates : pour les calculs sur les textes et dates.

Constantes et structures de contrôle : pour définir les algorithmes.



Menu Analyser

Pour analyser les tableaux à plat ou croisés : détailler le contenu des cellules, réaménager le tableau, choisir des graphiques, faire des tests statistiques.

Résultats

Choisir le mode de présentation

Pour voir les tableaux à plat, croisés ou de groupe sous forme de tableaux ou de graphiques.

Les **Tableaux récapitulatifs** donnent, par type de question, une synthèse des résultats.

Analyser le tableau : pour modifier la présentation du tableau, effectuer des tests statistiques, ou détailler les cellules (par un double clic sur une cellule du tableau).

Changer de strate

Définir une nouvelle strate ou choisir celle sur laquelle réaliser les analyses.

Le nom de la strate figure dans l'intitulé des fenêtres d'analyse.

Ok mémorise la strate et l'ajoute à la liste des strates.

Appliquer : lance le traitement courant sur cette strate.

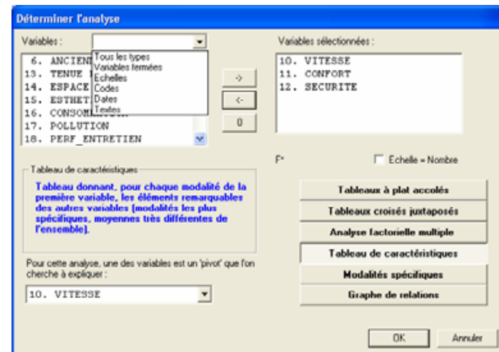
Déterminer une analyse (*)

Sélectionner à gauche un ensemble de variables et les faire passer dans la liste de droite.

Le logiciel donne la liste des traitements qui peuvent être effectués compte tenu du type des variables sélectionnées.

Cliquer pour voir les caractéristiques du traitement, et double cliquer pour le lancer.

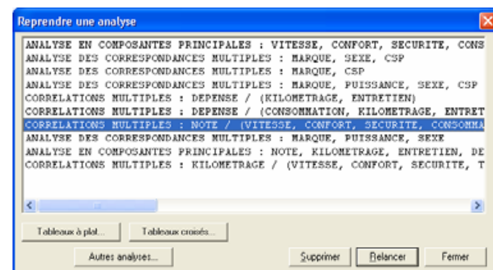
Pour un complément d'information sur les traitements, lancer les **Conseils pour l'analyse** depuis le panneau de commandes central.



Reprendre une analyse (*)

Pour relancer une analyse effectuée précédemment.

Faire un double clic dans la liste. Le traitement est relancé.



Tableaux à plat

Sélectionner la variable à analyser.

Les boutons permettent d'adapter la présentation :

- **choix d'un graphique** (**Graphique**).
- réaménagement des lignes (**Supprimer, Regrouper, Ordonner**) en les sélectionnant avec les triangles.
- **calculs des pourcentages** : par rapport aux observations, aux citations, en cumul (**Valeurs**).
- **tests statistiques** : intervalle de confiance, chi2 par rapport à l'équi-répartition ou à une référence (**Tests**), comparaison de fréquences (**Comparer**).

-**Recoder** pour conserver les modifications.

Le Sphinx Lexica - [automobiles - Traitements et analyses]

Fichier Edition Stade Dépouiller Recoder Analyser Approfondir ?

TABLEAU A PLAT : MARQUE - Echantillon total (402 observations)

Tableau | Graphique | Valeurs | Comparer | Ordonner | Regrouper | Renommer | Légende | Recoder | Tests | Caract. | Permuter | Supprimer | Revenir | Dimensions

MARQUE Quelle est la marque de votre voiture ?

MARQUE	Nb. cit.	Fréq.
Françaises	198	57,3%
Fiat	26	7,7%
Ford	26	7,6%
Toyota	5	1,5%
BMW	14	4,1%
Mercedes	12	3,7%
Opel	18	5,3%
Volvo	7	2,1%
Autre	36	10,7%
TOTAL CIT.	346	100%

Les valeurs sont pondérées par la variable 'REDRESSEMENT'.

Tableaux croisés

Sélectionner les variables à analyser. Les variables numériques sont automatiquement mises en classes, mais il est préférable de les analyser directement (corrélation).

Comme pour les tableaux à plat, les boutons permettent d'adapter la présentation :

- **choix d'un graphique**. **AFC** pour afficher une carte factorielle.
- **réaménagement** des lignes et des colonnes.
- **calculs** des pourcentages : en ligne, en colonne, sur le total, … (**Valeurs**).
- **tests statistiques** : test du chi2, C de Cramer, affichage des écarts significatifs (**Tests**), comparaison de distribution (**Comparer**).
- ajouter un commentaire (**Légende**).

Le Sphinx Lexica - [automobiles - Traitements et analyses]

TABLEAU CROISE : CSP ==> POSSESSION - Echantillon total (402 observations)

POSSESSION / CSP	Oui	Non	TOTAL
Non réponse	100	0,0	100
Agriculteur	78,7	21,3	100
Commerçant, artisan	100	0,0	100
Cadre	100	0,0	100
Employé	100	0,0	100
Ouvrier	89,3	10,7	100
Chomeur	85,7	14,3	100
Etudiant	68,5	31,5	100
Inactif	56,0	44,0	100
Autre	100	0,0	100
TOTAL	85,6	14,4	100

La dépendance est très significative. chi2 = 96, ddl = 9, 1-p = >99%.
Les cases encadrées en bleu (rose) sont celles pour lesquelles l'effectif réel est nettement supérieur (inférieur) à l'effectif théorique.
Attention, 5 (25.0%) cases ont un effectif théorique inférieur à 5, les règles du chi2 ne sont pas réellement applicables.
% de variance expliquée (V de Cramer) : 23,88%

Les valeurs du tableau sont les pourcentages en ligne établis sur 404 citations.
Les valeurs sont pondérées par la variable 'REDRESSEMENT'.

Tableaux de moyennes

Sélectionner la variable fermée puis les variables numériques ou échelles.

Comme pour les autres tableaux, les boutons permettent d'adapter la présentation : graphiques, réaménagement de lignes et colonnes…

- **choix d'un graphique** : **ACP** pour afficher une carte factorielle, **Dispersion** par critère ou modalité pour visualiser la moyenne et l'écart-type.
- **calculs** : le calcul de la moyenne peut être complété par l'écart-type, la somme, la part (% des sommes par catégories/somme totale) ou l'effectif (**Valeurs**).
- **tests statistiques** : analyse de la variance avec le test de Fisher (**Tests**).

Le Sphinx Lexica - [automobiles - Traitements et analyses]

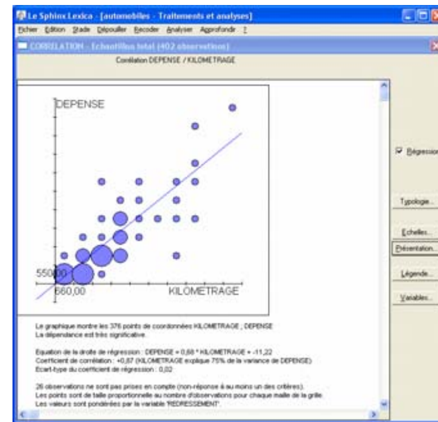
MOYENNES CROISEES DE MARQUE - Echantillon total (402 observations)

MARQUE	ENTRETIEN	KILOMETRAGE	DEPENSE
Non réponse	0	723	0
Renault	1	1480	925
Peugeot	1	1678	995
Citroën	1	1383	712
Volkswagen	2	1751	1047
Fiat	2	1824	1111
Ford	1	1213	803
Toyota	4	1050	971
BMW	3	2426	1948
Mercedes	3	1908	1663
Opel	2	1265	894
Volvo	2	818	559
Autre	2	1547	1211
TOTAL	1	1440	941

Les valeurs du tableau sont les moyennes calculées sans tenir compte des non-réponses.
Les noms des critères discriminants sont encadrés.
Les nombres encadrés correspondent à des moyennes par catégorie significativement différentes (test t) de l'ensemble de l'échan
Résultats du test de Fisher :
ENTRETIEN : V_inter = 20, V_intra = 5, F = 4, 1-p = >99%
KILOMETRAGE : V_inter = 5238066, V_intra = 790903, F = 7, 1-p = >99%
DEPENSE : V_inter = 4544139, V_intra = 449330, F = 10, 1-p = >99%

Corrélations

Sélectionner 2 variables numériques. Le résultat apparaît sous forme de graphe.



Présentation : Chaque point correspond à une observation. Visualiser les éventuelles superpositions par des points proportionnels aux effectifs (**Présentation**). L'équation de la droite de régression et le coefficient de corrélation apparaissent dans la légende.

Typologie : Pour définir des groupes de points par proximité sur le graphique. Choisir le nombre de classes désiré puis les positionner de manière à définir les groupes.

Variables : pour enregistrer la typologie.

Caractériser les cellules d'un tableau

Quel que soit le tableau étudié, un double clic sur une cellule donne accès aux caractéristiques des observations correspondantes :

Caractériser des observations

Les numéros des observations concernées
La variable 3. MARQUE
Le dépouillement de ces seules observations
Observations
Répartition
Spécificités
Graphique différentiel
Options...

Observations telles que OPTIONS = 'Jantes alu' (62 n°)
Spécificités pour la variable 'MARQUE' (Rapport des fréquences) :

	LOCAL	GLOBAL	INDICATEUR
Ford	16,1%	7,0%	+ (2,32)
Opel	9,7%	5,0%	+ (1,95)
Autre	16,1%	10,0%	+ (1,62)
Volvo	3,2%	2,0%	+ (1,62)
Peugeot	6,5%	10,4%	- (1,62)
Citroën	12,9%	10,0%	+ (1,30)
Mercedes	3,2%	4,0%	- (1,23)
Renault	16,1%	17,9%	- (1,11)
Volkswagen	9,7%	9,0%	+ (1,08)
Fiat	6,5%	6,5%	- (1,00)
BMW	0,0%	4,5%	(*)
Toyota	0,0%	1,5%	(*)

Transférer | Décrire... | Corriger... | Détailler... | Fermer

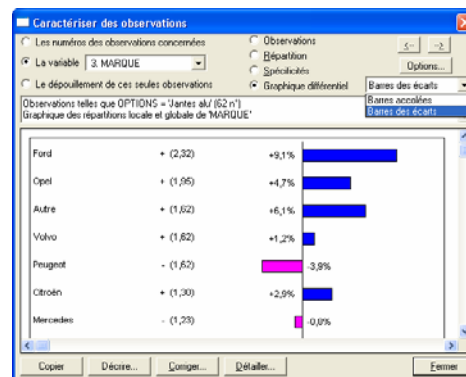
Observations : donne la liste des observations et leur valeur pour la variable sélectionnée.

Répartition : donne la répartition en pourcentage des modalités de la variable ou sa valeur moyenne et son écart-type.

Spécificité :

- comparaison de la répartition pour les effectifs de la cellule du tableau (% local) à l'effectif total (% global).

- indice de spécificité, rapport des fréquences ou test de significativité des écarts local global (contribution au chi2 ou test de Student).



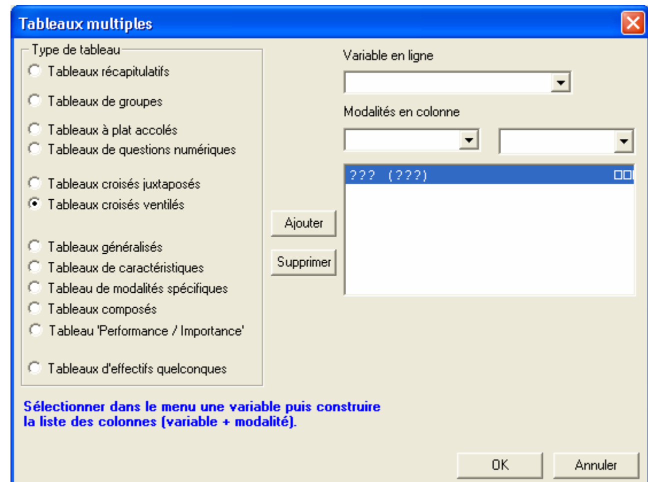
Graphique différentiel : au choix, graphique en barres de % globaux et locaux ou des écarts.

Dépouillement de ces seules variables : aperçu du dépouillement des effectifs concernés pour toutes les variables.

Corriger : pour passer en revue toutes les observations concernées et modifier certaines de leurs valeurs.

Tableaux multiples

Ces tableaux sont construits à partir de plusieurs variables combinées selon différents objectifs :

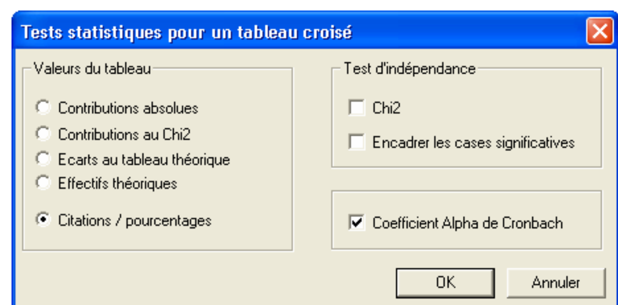


Résumé rapide : **Tableaux récapitulatifs**

Synthèses dans un même tableau de plusieurs tableaux à plat ou croisés (**Tableaux de groupes, Tableaux à plat accolés ou de questions numériques, Tableaux croisés juxtaposés ou ventilés**).

Analyse sélective en vue de faire ressortir les traits remarquables d'un grand tableau (**Tableaux de caractéristiques**, **Tableaux de modalités spécifiques**) ou d'une analyse complexe (**Tableaux composés**, **Tableaux importance/performance**).

Tableau d'effectifs quelconques : Entrée au clavier d'un tableau à analyser avec les procédures Sphinx.



Il est possible d'ajouter au menu **Analyser**, des articles pour les tableaux correspondant à une analyse sélective ou pour entrer un tableau au clavier.

Tableaux de groupes :

Pour analyser les groupes définis dans le questionnaire.

Tests statistiques : **Chi²**, **l'Alpha de Cronbach** pour évaluer la cohérence d'une batterie d'échelles.

Tableaux de caractéristiques

Pour faire ressortir les traits remarquables des modalités d'une variable fermée par rapport à plusieurs autres, fermées ou numériques. Ceci revient à réduire un grand tableau aux seules informations significatives.

Modalités les plus fréquentes : la sélection se fait en n'affichant que les caractéristiques les plus fréquentes ou la valeur moyenne.

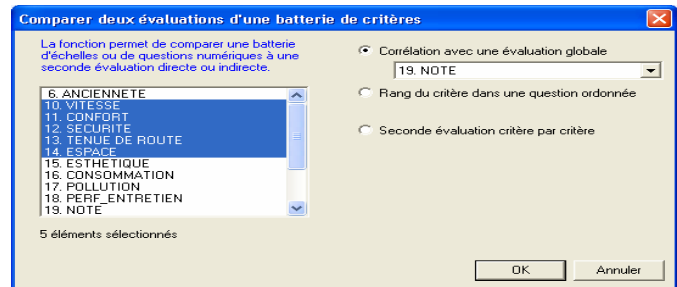
Modalités spécifiques : la sélection se fait selon les indications de tests statistiques.

Pour fixer les conditions selon lesquelles une information est jugée significative (bouton **Tests**) choisir entre :

- seuil de spécificité, comparaison de fréquence ou contribution au Chi2 (variable fermée) et test de Student (variables numériques).

- effectifs et sens de l'écart (sur / sous représentation ou supérieur / inférieur à la moyenne).

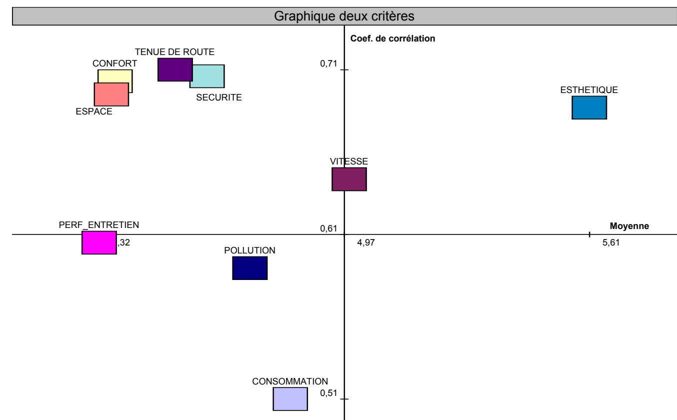
Tableau importance/performance



Corrélation à une évaluation globale

Pour représenter dans un même tableau les valeurs moyennes d'une batterie d'échelles (performance) et leur corrélation avec une évaluation globale (importance).

Ce tableau, représenté sous forme d'un graphique, fait ressortir les écarts entre importance et performance.



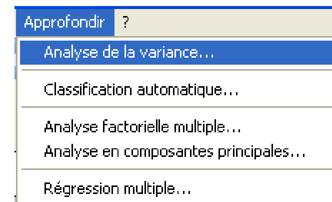
Rang du critère dans une question ordonnée

L'importance est déduite du rang donné aux critères dans une question à réponses ordonnées : classez par ordre décroissant l'importance à vos yeux des critères…

Seconde évaluation critère par critère

Comme l'importance, l'évaluation de la performance est faite sur une batterie d'échelles.

Menu Approfondir (*)



(*) Analyse multivariée :

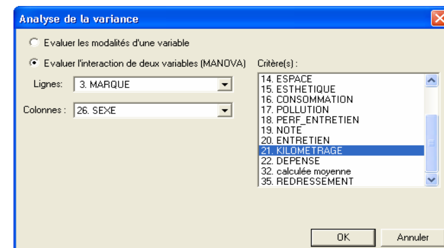
Traiter simultanément plusieurs variables pour :

- réduire la complexité des données (**AFM**, **ACP**)

-rechercher les influences multiples (**Analyse de la variance**, **régression multiple**),

-construire des typologies (**classification automatique**, **AFM**, **ACP**).

Analyse de la variance (MANOVA) (*)



Pour rechercher l'influence des catégories définies par le croisement de 2 variables nominales avec les valeurs d'une numérique.

Sélectionner les 2 variables nominales explicatives puis la variable numérique à expliquer.

Présentation du tableau :

Suppression, regroupement, classement des lignes et/ou colonnes.

Choix des valeurs : moyenne, écart-type, somme, part, effectifs (**Valeurs**).

Tests statistiques

Analyse de la variance à un facteur.

Analyse de la variance à 2 facteurs : donne l'effet principal de chaque variable et de l'interaction.

Le Sphinx Lexica - [automobiles - Traitements et analyses]

Fichier Edition Stade Dépouiller Recoder Analyser Approfondir ?

TABLEAU DE MOYENNES - Echantillon total (402 observations)

Tableau Graphique Valeurs Comparer Ordonner Regrouper Renommer Légende Carrousel
Tests Caract. Permuter Supprimer Revenir Dimensions Transposer

	Homme	Femme	TOTAL
Renault	2162,50	1235,00	1647,22
Peugeot	1625,00	1700,00	1670,73
Citroën	1560,00	1304,00	1432,00
Volkswagen	2562,50	1460,00	1950,00
Fiat	1550,00	1928,57	1753,85
Toyota	2000,00	200,00	1400,00
BMW	2857,14	1750,00	2611,11
Mercedes	2583,33	1000,00	2187,50
Opel	1733,33	950,00	1420,00
Volvo	1250,00	500,00	1000,00
Autre	1616,67	1500,00	1570,00
TOTAL	1985,54	1407,61	1706,48

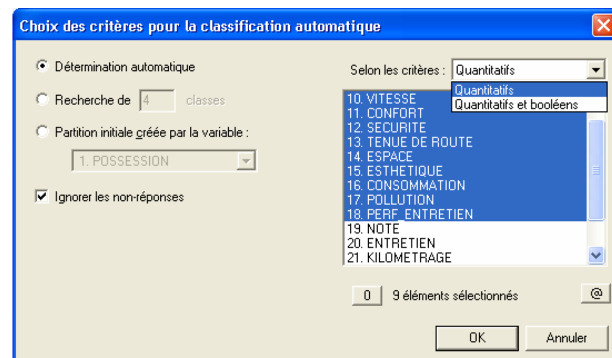
Les valeurs du tableau sont les moyennes de la variable KILOMETRAGE pour chaque couple de citations.

Analyse de la variance à un facteur :
- L'effet global de 'MARQUE' est très significatif (V_inter = 2988926,65, V_intra = 1001930,50, F = 29,83, 1-p = 99,86%)
- L'effet global de 'SEXE' est très significatif (V_inter = 6687889,03, V_intra = 983434,23, F = 6,80, 1-p = 99,07%)
- Les modalités discriminantes (au risque de 95%) sont encadrées

Analyse de la variance à deux facteurs :
- L'effet principal de 'MARQUE' est très significatif (V_inter = 2988926,65, V_intra = 863553,78, F = 3,46, 1-p = 99,97%)
- L'effet principal de 'SEXE' est très significatif (V_inter = 6687889,03, V_intra = 863553,78, F = 7,74, 1-p = 99,43%)
- L'interaction de 'MARQUE' et 'SEXE' est très significative (V_inter = 4052666,28, V_intra = 863553,78, F = 4,69, 1-p = >99,99%)

Classification automatique (*)

Pour regrouper les observations en classes homogènes relativement à un ensemble de critères quantitatifs et/ou booléens (recoder les fermées multiples en variables oui/non).

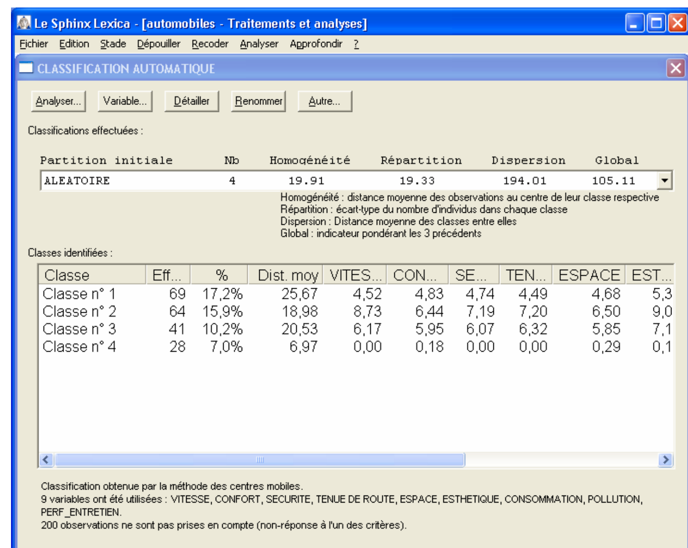


Détermination automatique Plusieurs classifications au nombre de classes croissant sont proposées. Le nombre de classes est directement fixe ou résulte du choix d'une partition initiale.

Liste des classifications

Dérouler la liste **partition initiale** pour sélectionner une classification selon son nombre de classes, son homogénéité, sa répartition et sa dispersion.

La composition de la classification est décrite dans la fenêtre **Classes identifiées**.



Détailler : pour voir les caractéristiques des observations de la classe sélectionnée.

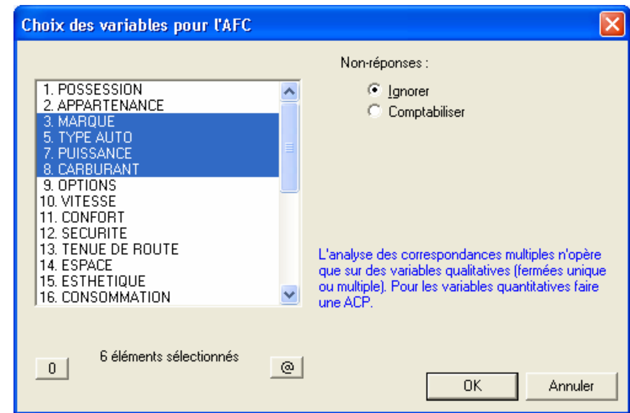
Renommer : pour nommer les classes en fonction de leurs caractéristiques.

Variable : pour créer une variable affectant chaque observation à la classe à laquelle elle appartient.

Analyser : pour lancer le tableau des valeurs moyennes des classes pour les critères, et l'analyser.

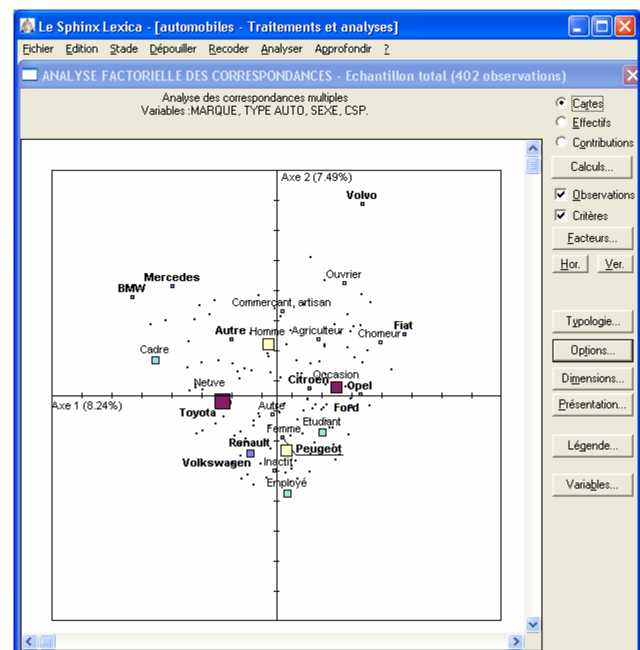
Analyse factorielle multiple (*)

Pour réduire le nombre des éléments décrivant les observations en cherchant à les ramener à quelques facteurs synthétiques et à construire une typologie.



L'analyse factorielle multiple s'applique aux données qualitatives.

Les résultats de l'analyse apparaissent sous forme de :



Cartes : projection sur le plan des axes factoriels (Facteurs…, Hor., Ver).

Effectifs : tableau des effectifs croisés.

Contributions : des modalités des variables aux axes factoriels.

Calculs : détail de tous les paramètres (valeurs propres, variance expliquée, coordonnées et contributions des modalités, cosinus2).

Présentation : les points individus sur la carte sont 'cliquables' pour voir leurs caractéristiques. La taille des points et symboles peut être modifiée et les libellés déplacés.

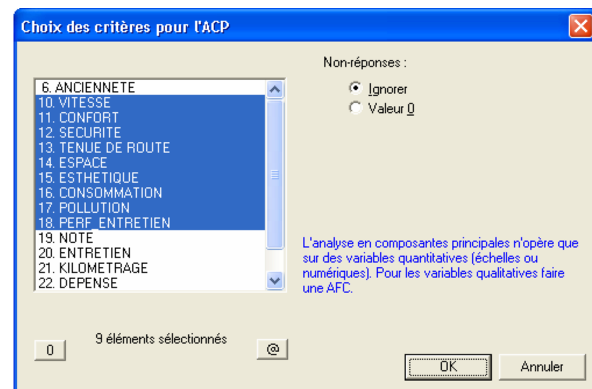
Typologie : pour définir par proximité des groupes de points sur le graphique. Choisir le nombre de classes désiré puis les positionner de manière à définir les groupes.

Variables : pour enregistrer la typologie.

Analyse en composantes principales (*)

Pour ramener plusieurs mesures à un plus petit nombre de dimensions et construire une typologie.

L'analyse en composantes principales s'applique aux données quantitatives, variables numériques ou échelles.



Les résultats de l'analyse apparaissent sous forme de :

Cartes : projection sur le plan des axes factoriels (Facteurs…, Hor., Ver.).

Corrélation : matrice des coefficients de corrélation entre les variables choisies.

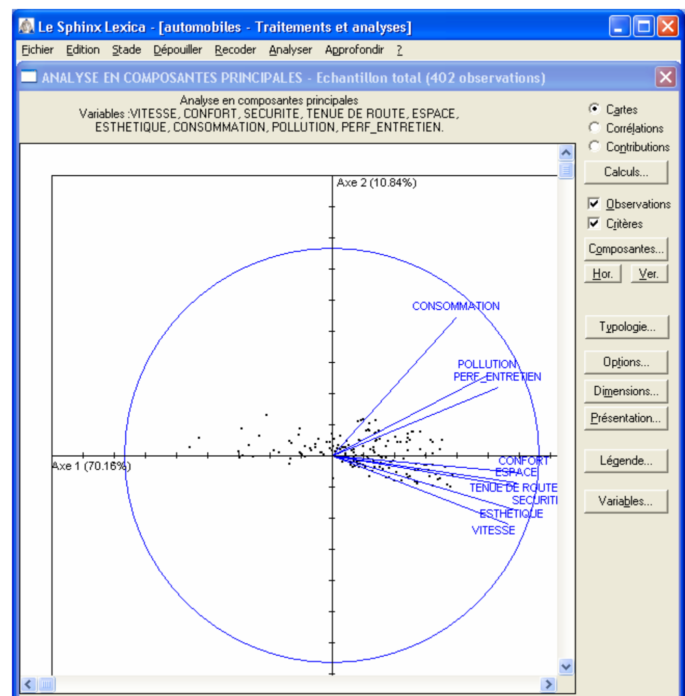
Contributions des variables aux composantes (axes factoriels).

Calculs : détail de tous les paramètres (valeurs propres, variance expliquée, coordonnées des variables et des individus et contribution des variables, cosinus2).

Présentation : les points individus sur la carte sont 'cliquables' pour voir leurs caractéristiques. La taille des points et symboles peut être modifiée et les libellés déplacés.

Typologie : pour définir par proximité des groupes de points sur le graphique. Choisir le nombre de classes désiré puis les positionner de manière à définir les groupes.

Variables : pour enregistrer la typologie.



Atelier lexical (*)

Les raccourcis (**)

Dans Lexica, le panneau de commandes pour l'étude des textes propose des raccourcis pour directement activer des procédures suivantes, appliquées à la variable sélectionnée.



- Créer les variables lexicales (texte lemmatisé, mesures lexicales...)
- Construction de lexiques
- Recherche des concordances d'un mot
- Recherche des segments répétés
- Bilan lexical par contexte
- Construction de lexiques croisés
- Recherche de termes spécifiques
- Réponses caractéristiques

Ces résultats peuvent être affinés dans l'atelier lexical.

Navigation lexicale (*)

A gauche le lexique, à droite les réponses.

Réduire : pour enlever les mots outils

Regrouper, **Supprimer** : pour agir sur les mots sélectionnés dans le lexique.

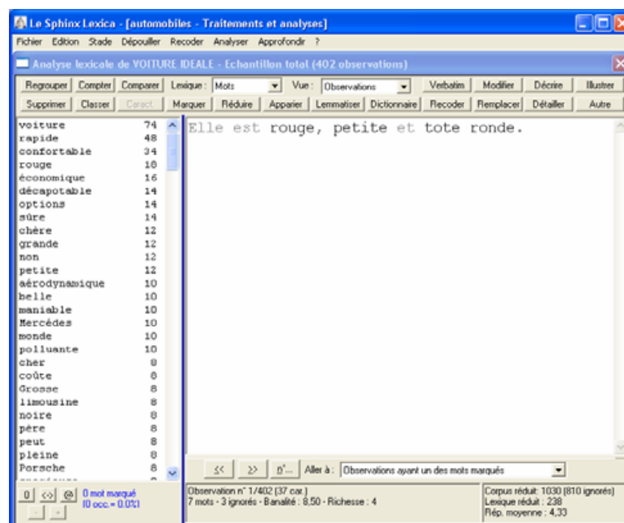
Marquer : pour chercher des mots.

Flèches de navigation : pour trouver dans le texte les réponses contenant le ou les mots marqués.

Aller à : pour paramétrer la navigation (un mot marqué, tous les mots marqués…).

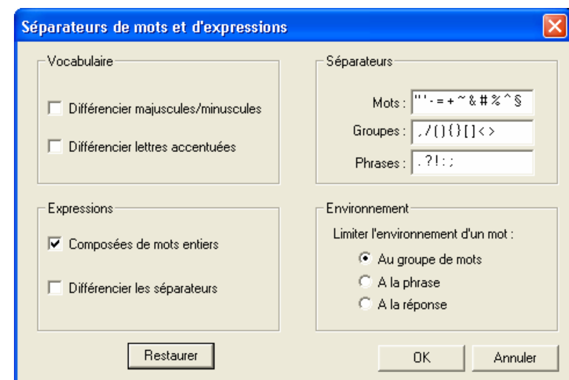
Double clic sur un mot : pour voir la liste des réponses contenant le mot (concordances).

Verbatim : pour sortir les résultats.

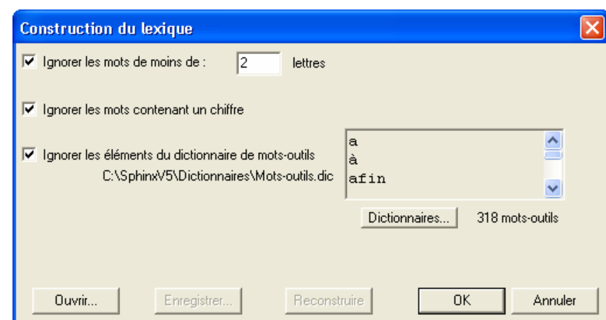


Production des lexiques (*)

Apparier : ce bouton (2ème ligne au centre) définit les séparateurs en fonction desquels le lexique est calculé en identifiant les mots comme séquence de caractères sans séparateurs.



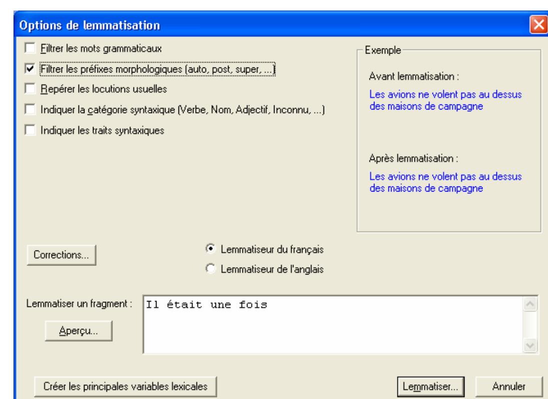
Réduire : pour éliminer du calcul du lexique certaines formes graphiques (selon le nombre de caractères, la présence de chiffres ou l'appartenance au dictionnaire, des mots autres).



Lemmatisation (**)

Lemmatiser (2ème rang au centre de la barre d'outils) : pour lancer l'analyse syntaxique et créer une version du corpus réduite aux formes racines du texte, (masculin singulier des noms adjectifs et pronoms, infinitif des verbes).

Créer les principales variables lexicales : pour lancer en une seule fois la création des versions lemmatisées du corpus et/ou réduites aux noms, verbes, adjectifs, ainsi que les mesures lexicales (longueur, richesse, banalité)…



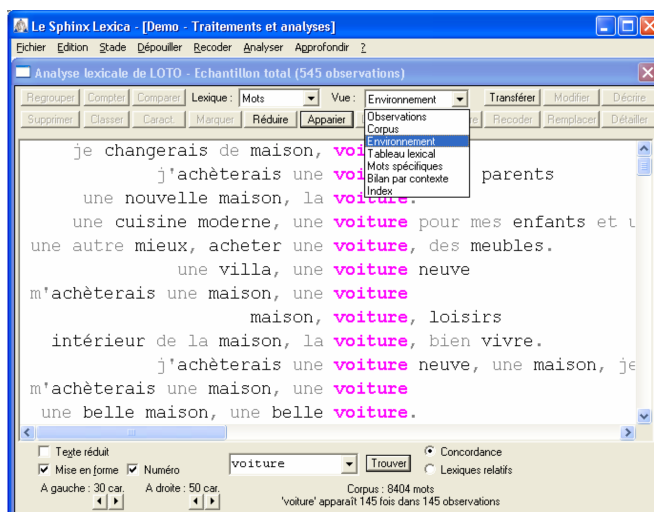
Environnement : Concordances (**)

Article **Environnement** du menu **Vue**, ou double clic sur un mot du lexique affiche l'environnement du mot :

- ensemble de la réponse,
- phrase,
- groupe de mots.

Contenant le mot sélectionné.

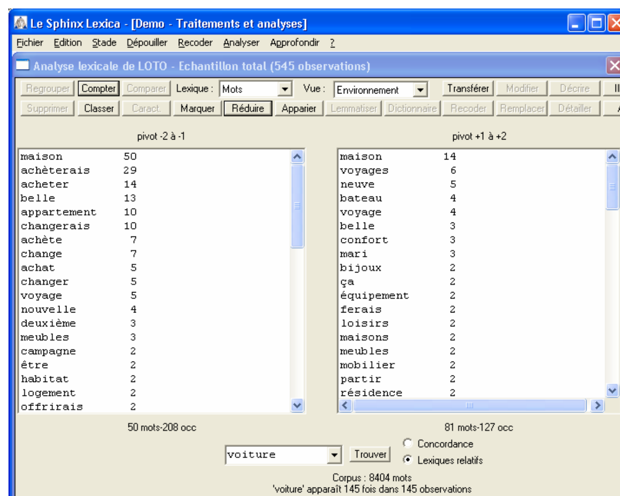
Appartier : pour définir l'environnement par le choix des séparateurs.



Environnement : Lexiques relatifs (**)

Lexique à gauche et à droite du mot pivot.

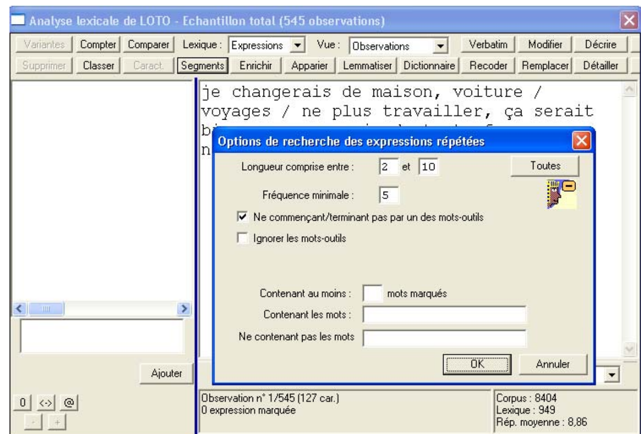
Compter : pour définir l'étendue de l'environnement à gauche et à droite, pour le calcul des lexiques.



Expressions et segments répétés (**)

Expressions : dans la liste déroulante **Lexique, Expressions** donne accès à la recherche des locutions entrées au clavier ou provenant d'un dictionnaire. (**Dictionnaire**).

Segments : pour spécifier la recherche des segments répétés du texte. Longueur des segments, mots présents dans le segment…



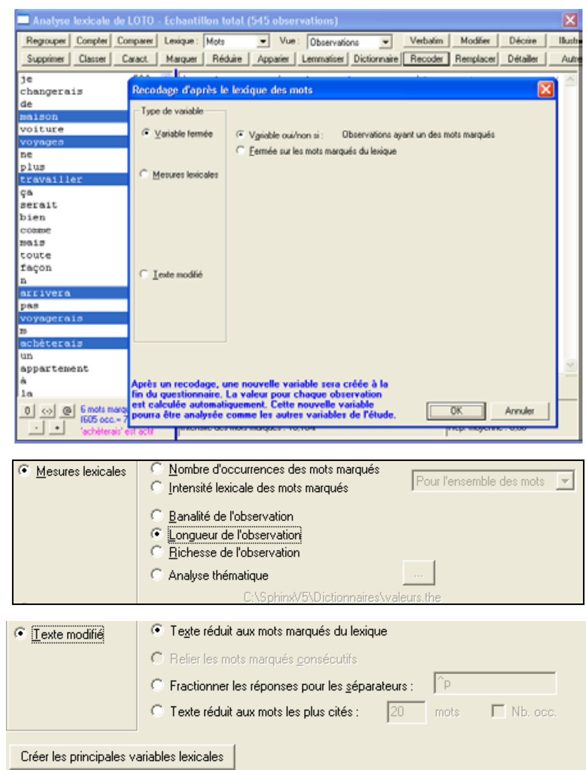
Recodifications (*)

Recoder : (2ème ligne sur la gauche de la barre d'outils) pour créer une nouvelle variable à partir des mots marqués dans le lexique.

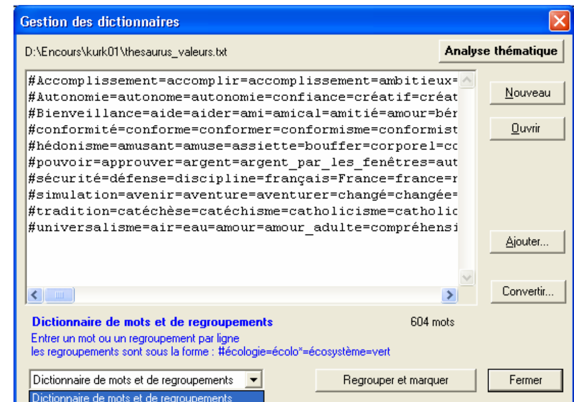
Variable fermée : une variable oui/ non pour chaque mots marqués ou une variable fermée multiple dont les modalités sont les mots marqués du lexique.

Mesures lexicales : variables numériques caractérisant les réponses (longueur, intensité lexicale, banalité…).

Texte modifié : pour relier les expressions et les traiter comme un mot ou transformer le texte en l'expurgeant ou en l'éclatant.



Dictionnaires (*)



Dictionnaires de mots : liste de mots placés dans le dictionnaire à partir de sélection dans le corpus (**Ajouter**) ou directement entrée au clavier.

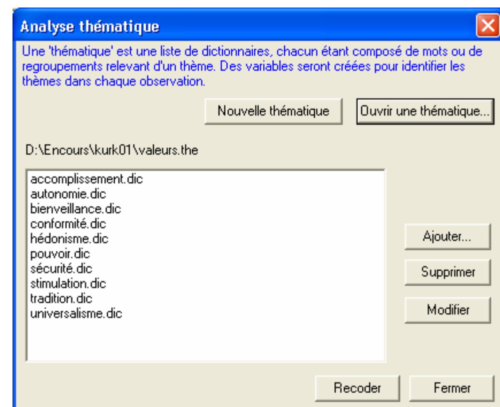
Marquer : pour repérer dans le lexique les mots du dictionnaire.

Dictionnaire de mots et regroupements : liste de mots équivalents portés sur une ligne et définissant une catégorie.

Convertir : pour remplacer dans le corpus l'ensemble des termes d'une ligne par le nom de la catégorie correspondante et ainsi regrouper les termes équivalents.

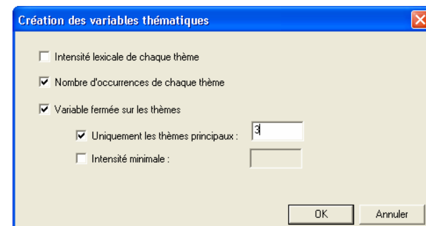
Analyse thématique (*)

Pour mobiliser en une seule fois plusieurs dictionnaires de mots ou de regroupements et créer des variables caractérisant la présence des éléments de ces dictionnaires dans le corpus : calcul de variables numériques (intensités lexicales ou nombre d'occurrences) pour thème ou d'une variable fermée dont les modalités correspondent aux thèmes.



Nouvelle thématique : pour définir une thématique en sélectionnant les dictionnaires qui la composent.

Ouvrir une thématique : pour reprendre une thématique préalablement créée.



Bilan par contexte (**)

Dans le menu **Vue** : choisir **Bilan par contexte.**

Donne les caractéristiques du texte selon les modalités d'une variable de contexte.

Croiser : pour choisir la variable de contexte.

Réduire : pour relancer le calcul sans les mots outils…

Tableau lexical (**)

Ventile les lexiques selon les modalités d'une variable de contexte.

Dans le menu **Vue** : choisir **Tableau lexical**.

Croiser : pour définir la variable de contexte.

Compter : pour calculer l'indice de spécificité indiquant si le mot est sur/sous représenté.

Mots spécifiques (**)

Donne pour chaque modalité de la variable de contexte la liste des mots qui lui sont spécifiques (sur représentés).

Dans le menu **Vue** : choisir **Mots spécifiques**

Croiser : pour définir la variable de contexte.

Sélect : pour indiquer comment filtrer les mots spécifiques :

- Selon que le mot est exclusif à la catégorie, présent dans toutes les catégories ou dans plusieurs.

- Selon la valeur de l'indice de spécificité et la fréquence.

- Selon un nombre total de mots à afficher.

Réponses caractéristiques : pour sortir les extraits correspondant à la plus forte intensité lexicale des mots spécifiques sélectionnés.

Index (**)

L'index donne la liste des numéros d'observations contenant les éléments sélectionnés dans le lexique.

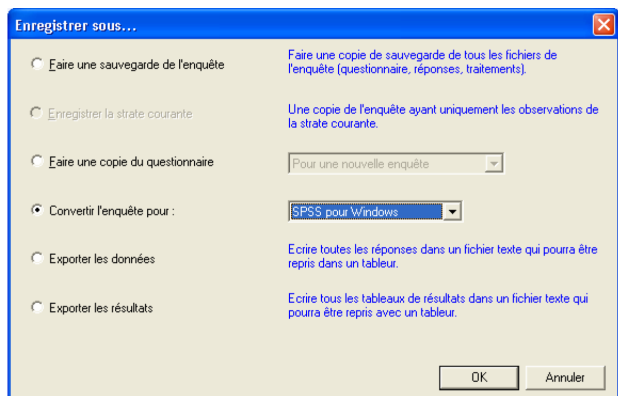
Menu Fichier

Pour ouvrir et enregistrer les applications Sphinx ou échanger avec d'autres applications.

Rappel de la syntaxe des fichiers sphinx :

- nom.que : le questionnaire.
- nom.rep : les réponses fermées et numériques.
- nom.ouv : les réponses aux questions ouvertes
- nom.hyp : les tableaux de bord et formulaires.

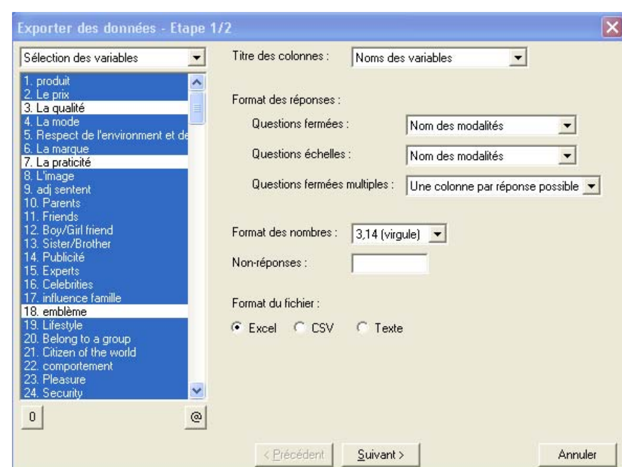
Faire une Sauvegarde / Convertir



Convertir l'enquête : pour analyser les données avec :

- SPSS.
- les logiciels statistiques compatibles avec le format Triple-s (SAS notamment).
- Alceste pour l'analyse des données textuelles.
- les versions antérieures de Sphinx.

Faire une sauvegarde / Exporter



Sélectionner les variables à exporter et définir le format d'exportation (**Excel, Csv, Texte**).

Pour les questions à réponses fermées multiples, le choix est offert entre plusieurs possibilités :

- Une colonne par réponse possible ;
- Une colonne par modalité ;
- Une seule colonne contenant le texte des modalités cochées.

Inclure dans le rapport

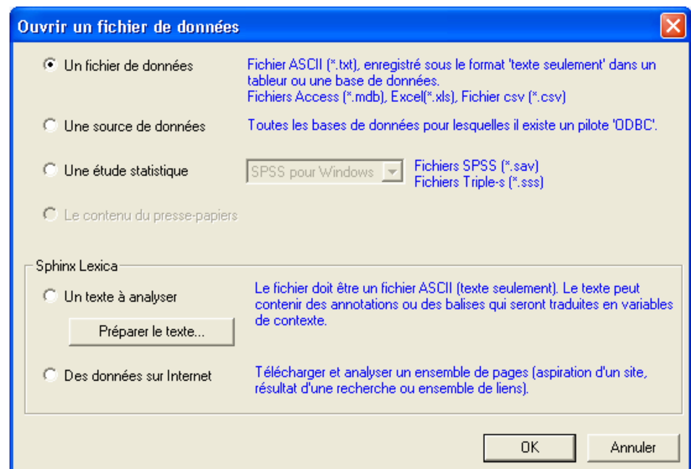


Transfert dans Word des résultats affichés à l'écran.

Image : les résultats sont recopiés comme image.

Texte : pour construire les tableaux au format du traitement de textes utilisé.

Ouvrir un fichier de données (*)



Pour importer des données externes en les convertissant au format Sphinx.

Sphinx reconnaît les fichiers .txt, les formats d'Excel, Access, Dbase, Spss, de toutes les sources de données compatibles ODBC et fichiers au standard Triple-s

Assistant d'importation de données (*)

1/ Indiquer le nom du fichier à ouvrir, vous obtenez un aperçu de son contenu.

2/ Définir les séparateurs de variables et d'observations (uniquement si le format importé n'est pas standard).

3/ Contrôler le type de variable affecté à chaque colonne. Eventuellement modifier les éléments : nom de variables, modalités, type.

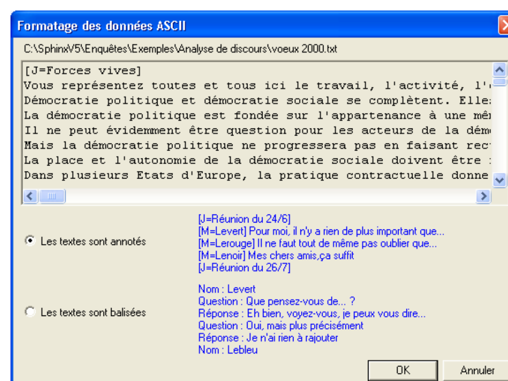
4/ Corriger les erreurs détectées (non-conformité entre le format et le contenu) en éditant directement les valeurs marquées en rouge.

Ouvrir un texte à analyser (**)

Choix du type de structuration (**)

Pour indiquer comment le texte importé est structuré.

Le logiciel exploite les marques et balise du texte pour le découper en fragments et construire une base Sphinx.

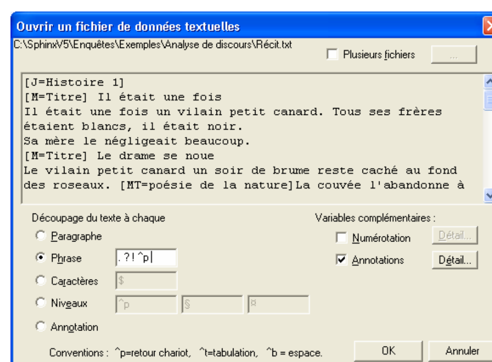


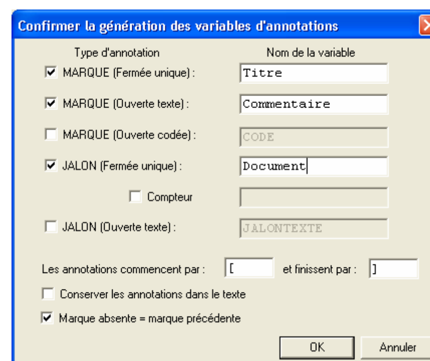
Texte annoté : définir la structuration (**)

Texte annoté : des annotations portées dans le texte indiquent la nature du fragment dans lequel elles ont été insérées. Les fragments sont définis par les règles de découpage (**phrase**, **paragraphe**, ou autre **caractère**….).

Les marques s'appliquent à un seul fragment considéré, les jalons portent sur tous les fragments jusqu'au prochain jalon.

Détail : pour nommer les variables correspondant aux Marques et Jalons.

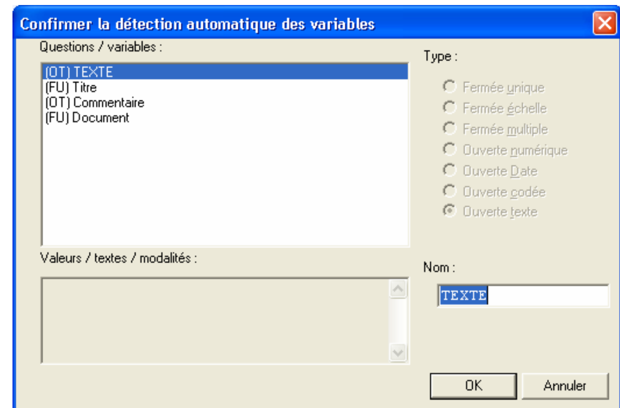




Texte annoté : création des variables (**)

Le texte est converti dans une base Sphinx dont les variables contiennent le texte et les informations de contexte données par les marques et les jalons.

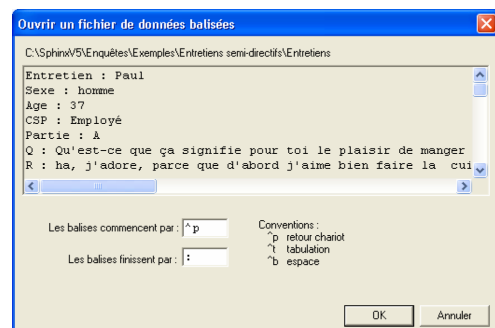
Le nom et le type de ces variables peuvent être modifiés avant la création de la base.



Texte balisé : définir la structuration (**)

Texte balisé : chaque élément du texte est défini par une balise.

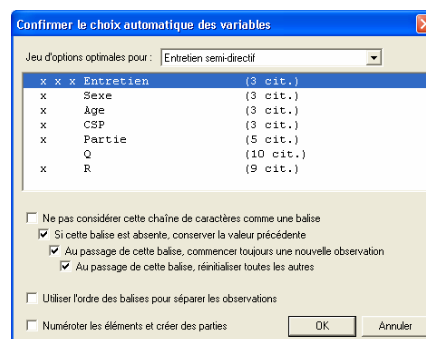
Indiquer les caractères qui délimitent les balises. Attention, si ces caractères figurent ailleurs dans le texte, la reconnaissance de la structure peut être perturbée.



Texte balisé : définition des variables (**)

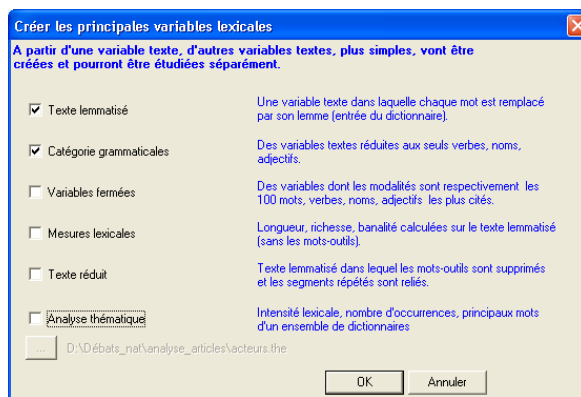
Le texte est converti dans une base Sphinx dont les variables correspondent aux balises. Une nouvelle observation est créée lorsqu'une balise apparaît à nouveau. Les balises absentes sont mises à non réponse ou reprennent la dernière valeur rencontrée.

Le réglage de la reconnaissance des balises permet ainsi de s'adapter au type de corpus traité (interview non directives, bases de données bibliographiques…).



Génération de l'enquête et création des variables lexicales (**)

A la création de la base Sphinx, le texte peut être soumis à l'analyse syntaxique et à un premier traitement statistique pour créer automatiquement des variables contenant le texte lemmatisé, le texte réduit aux noms, verbes, adjectifs, les principales mesures lexicales (longueur, richesse, banalité…).



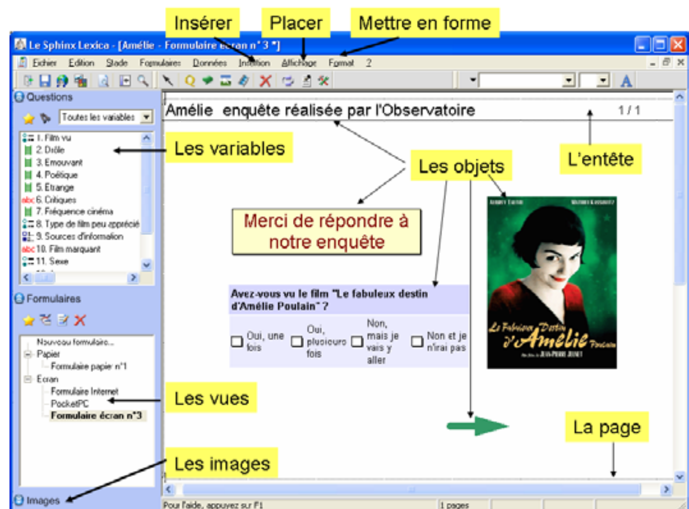
12. Formulaires multimédia

L'environnement de travail

L'environnement de travail multimédia permet de :

- Placer dans un formulaire les questions accessibles dans la liste des variables.
- Créer plusieurs formulaires répertoriés dans la liste des formulaires.
- Agrémenter les formulaires de légendes, et illustrations.

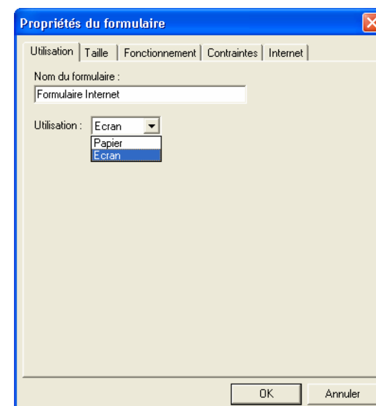
Les objets (questions, légendes, illustrations) sont placés dans la vue (formulaire) en utilisant les menus et raccourcis de la barre d'outils.



Propriétés du formulaire

Utilisation

Pour faire apparaître le panneau de propriétés des formulaires : **Fichier / Mise en page** ou bien, cliquer avec le bouton droit dans la vue.



Nom du formulaire :

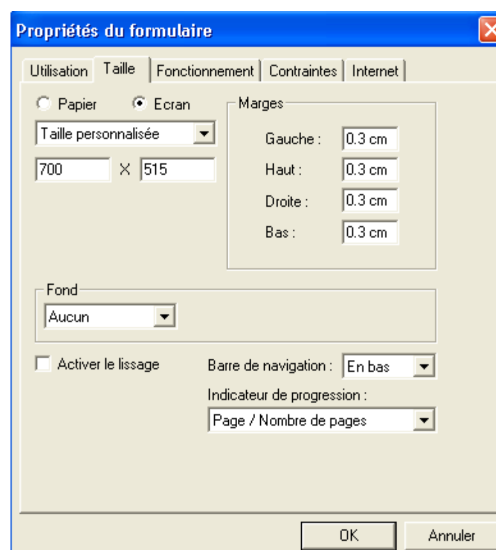
Pour concevoir un formulaire adapté à chaque usage, ou correspondant à différentes versions (interrogation complète ou limitée à quelques questions…).

Utilisation :

Type d'utilisation du formulaire sur **Papier** ou **Ecran**.

Taille

L'onglet **Taille** du panneau de propriétés des formulaires permet de choisir la taille de la vue selon l'utilisation et le fond (couleur, papier peint, illustration).



Barre de navigation :

Pour choisir la position des flèches de passage à l'écran suivant (**En bas** ou **En haut**).

Choisir l'option **Aucune** si la barre de navigation est définie directement dans le formulaire par des boutons personnalisés.

L'option **Choisir comme masque** du menu contextuel permet d'insérer ces boutons sur toutes les pages du formulaire.

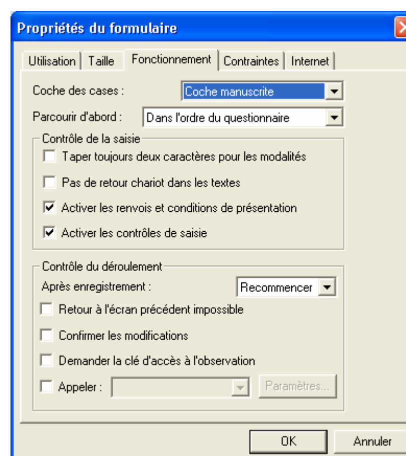
Activer le lissage :

Le lissage a pour effet de rendre les couleurs transparentes et de permettre la lecture des éléments superposés.

Indicateur de progression :

Pour faire apparaître la progression de l'enquête selon le nombre de pages ou avec une barre de progression.

Fonctionnement



Coches des cases :

Pour déterminer le style des coches : soulignées, normales, manuscrites.

Parcourir d'abord :

Pour définir l'ordre de passage d'une question à l'autre au moment de la saisie (ordre des questions, haut en bas, gauche à droite).

Contrôle de la saisie :

Pour régler le travail de l'opérateur de saisie.

Contrôle du déroulement :

Règle le passage d'un écran à l'autre et fixe le dialogue après enregistrement.

Contraintes

Permet de vérifier que les réponses sont cohérentes et ne violent pas les conditions fixées par les règles d'intégrité.

On ajoute autant de conditions que nécessaire.

Ci-dessous la fenêtre de définition du profil spécifiant une contrainte.

Fonctionnement Internet

Transactions en saisie :

S'il n'y a pas de renvois ou d'affichages conditionnels, on accélère le passage d'un écran à l'autre en choisissant : **Jamais**.

Format des images :

Pour forcer l'affichage des images dans un format donné (.GIF - .JPEG .PNG). Attention aux images trop lourdes qui ralentissent l'affichage.

Propriétés des questions

Contenu

On coche dans l'onglet **Contenu** les éléments qui apparaîtront sur le formulaire :

Titre de question :

Le libellé et le numéro de la question, ainsi que le nom de variable.

Largeur fixe :

Permet de standardiser la largeur de la zone Libellé.

Zone de saisie :

Pour afficher la zone destinée à entrer la réponse.

Consigne :

Pour ajouter une zone éditable.

Multimédia :

Pour illustrer la question avec une image.

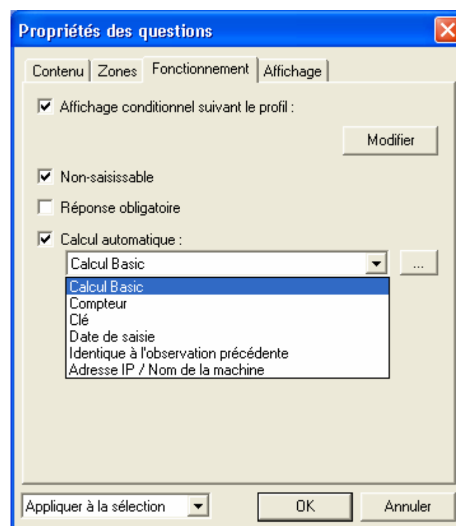
Disposition des zones de saisie :

Variable :

Changer le texte de la question ou modifier son contenu (libellé et modalité) en cliquant sur **...**

Dans le cas d'un groupe de questions, la liste des variables du groupe peut être directement modifiée.

Fonctionnement



Affichage conditionnel :

La question n'apparaît que si les conditions définies par profil sont remplies. C'est une alternative à la définition d'une condition dans le questionnaire.

Non saisissable :

Pour afficher des informations sans permettre de les modifier.

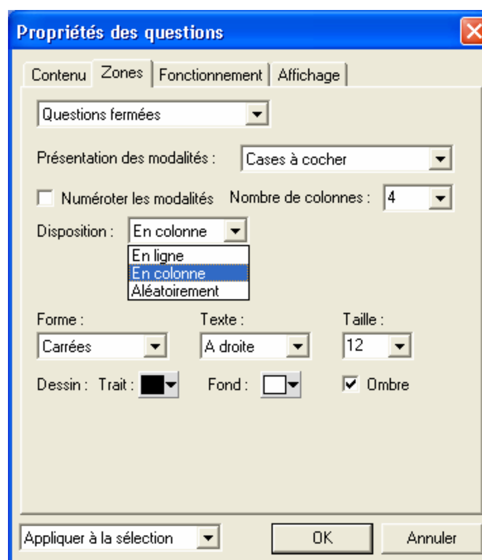
Calcul automatique :

La réponse n'a pas à être saisie.

Identique à l'observation précédente :

Pour faciliter la saisie d'informations répétitives – l'identité de la personne faisant la saisie par exemple.

Zones : questions fermées



Présentation des modalités :

Pour sélectionner le mode de présentation des modalités : **Cases à cocher - Modalités encadrées – Pictogrammes - Menu déroulant – Liste - Echelle graduée.**

La forme des coches est fixée dans les propriétés du formulaire, dans l'onglet **Fonctionnement**.

Forme – Texte – Taille :

Pour jouer sur la police, la forme, la couleur la disposition des cases à cocher.

Disposition :

Permet de fixer le mode d'affichage des modalités de réponses selon l'ordre dans lequel elles ont été créées (**En ligne** ou **En colonne**). **Affichage aléatoire** permet d'afficher, à chaque consultation, les modalités dans un ordre différent fixé d'une manière aléatoire. On élimine ainsi les effets d'ancrage.

Zones : questions fermées ordonnées

Le répondant clique selon l'ordre de ses préférences

Dans la liste suivante, quels sont pour vous les cinq principaux points à prendre en considération pour l'achat d'une automobile ?

La vitesse	1 La consommation	Le coût d'entretien
Le confort	La tenue de route	2 L'espace
La sécurité	3 La pollution	L'esthétique

Propriétés des questions
Contenu | Zones | Fonctionnement | Affichage
Questions fermées ordonnées
Présentation des modalités : Pictogrammes
Numéroter les modalités Nombre de colonnes : 4
Disposition : Aléatoirement
Pictogrammes :
Pictos_ecran\Etoiles1.gif;Pictos_ecran\Etoiles2.gif ...
Texte :
A droite
Dessin : Trait : Fond : Ombre
Appliquer à la sélection OK Annuler

Pictogrammes :

Des exemples de pictogrammes sont accessibles avec le bouton **...**

Toute image peut être associée à une modalité. Les fichiers correspondants sont séparés par des ;

La liste déroulante permet de choisir un ensemble prédéfini de pictogrammes. Le bouton **...** permet de choisir ses propres pictogrammes.

Attention : ne pas confondre avec l'illustration de la question (onglet contenu, multimédia).

Zones : questions ouvertes

Présentation du texte :

Zone vierge, lignes, cadre délimitant la zone, cases pour les caractères, hirondelles.

Contenu | Zones | Fonctionnement | Affichage
Questions numériques
Présentation du texte : Cases pour les caractères
Nombre de chiffres : 5

Zones : questions dates et codes

Date : Le format peut contenir jour, mois, année, heure, minute, seconde. Par exemple :

mm/aaaa pour le mois et l'année sur 4 chiffres ou

hh/mm/ss pour les heures, minutes et secondes…

Code : On associe une liste pour :

- contrôler la saisie : seuls les codes apparaissant dans le fichier pourront être saisis.

- interpréter le code : dans les tableaux de bord, on pourra regrouper les codes en utilisant les catégories définies dans cette liste.

Par exemple pour les départements :

Code=Département=Région=Académie=Nord/sud=IDinsee

01=Ain=RHONE ALPES=Lyon=SUD=82

02=Aisne=PICARDIE=Amiens=NORD=22

03=Allier=AUVERGNE=Clermont-Ferrand=NORD=83

04=Alpes-de-Haute-Provence=PROVENCE ALPES COTE D'AZUR=Aix-Marseille=SUD=93

05=Hautes-Alpes=PROVENCE ALPES COTE D'AZUR=Aix-Marseille=SUD=93

Affichage

Mise en forme des éléments :

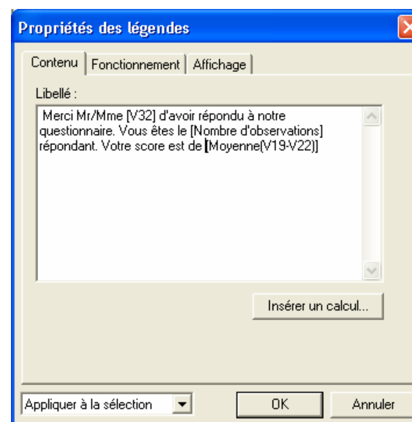
Pour choisir une mise en forme adaptée (polices, couleurs, alignement) pour chaque élément de la question : libellé (texte de la question), saisie (modalités de réponse), réponses (coche ou texte de la réponse), consigne.

Couleur de fond :

En choisissant 2 couleurs pour la zone de saisie on fait alterner les couleurs de chaque ligne pour faciliter la lecture des tableaux.

Propriétés des légendes

Contenu

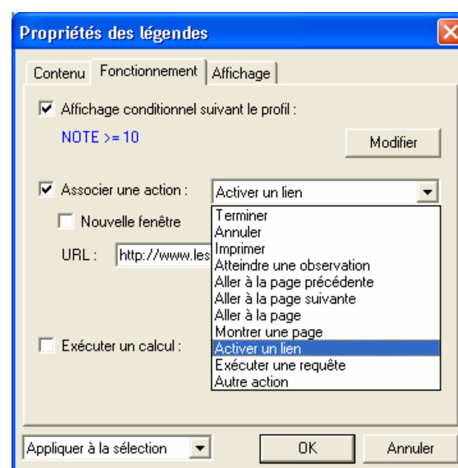


Le texte de la légende peut contenir les valeurs d'une variable. Ci-contre, on note [V32] pour afficher le contenu de la variable 32. On aurait pu aussi écrire [Nom]. Les crochets indiquent que la valeur correspondante doit être affichée dans la légende.

Insérer un calcul :

Cette valeur peut être ce que renvoie une fonction : [Nombre d'observations] ou le résultat d'un calcul : [moyenne (« V19-V22 »).

Fonctionnement



Affichage conditionnel :

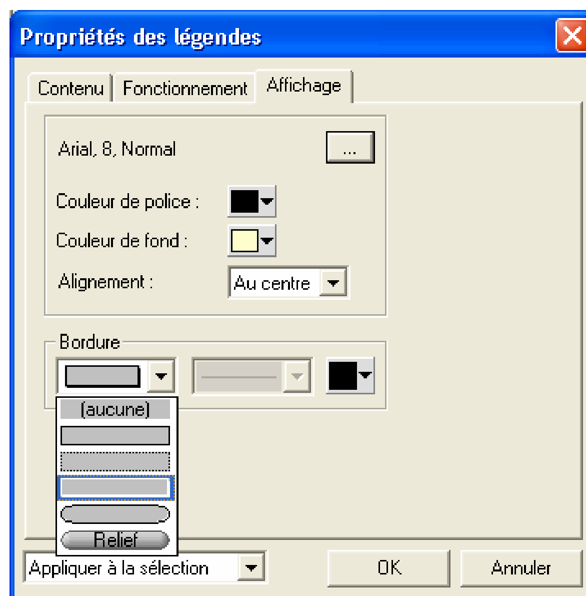
L'affichage de la légende peut être conditionné par un profil.

Associer une action :

La légende peut renvoyer de l'information disponible sur un site ou renvoyer à une autre page (c'est aussi une manière de faire des renvois), ou provoquer l'enregistrement des réponses avec l'action **Terminer**.

Affichage

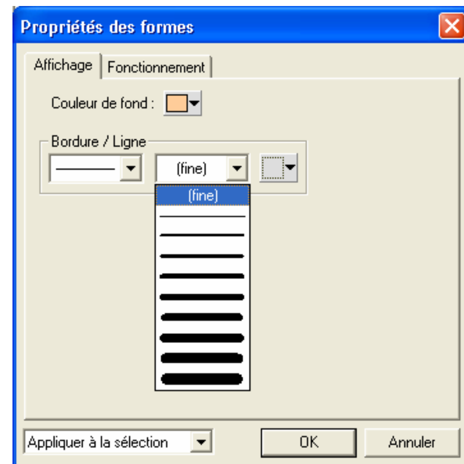
Pour mettre en forme la légende : police, couleur de fond, encadrement…



Propriétés des formes

Affichage

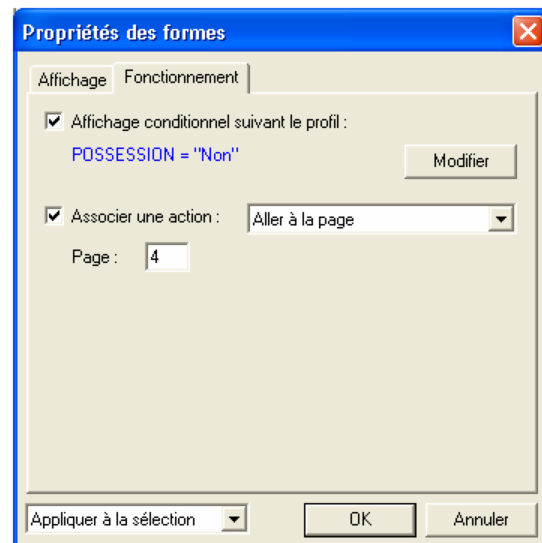
Les formes graphiques peuvent être utilisées pour illustrer ou délimiter un formulaire…



Fonctionnement

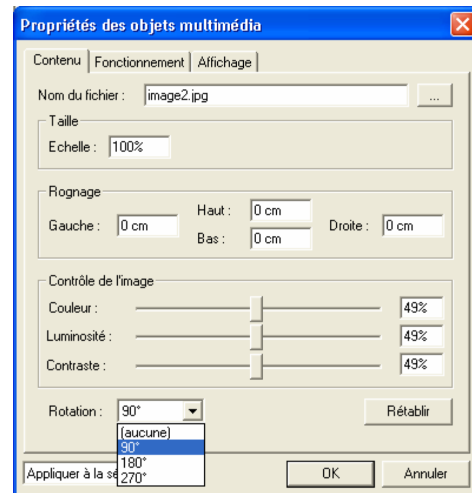
Les formes fonctionnent comme les autres objets.

On peut ainsi organiser des sauts en superposant des formes flèches qui envoient à des pages différentes selon la réponse à une question filtre et élaborer ainsi des scénarios.



Propriétés des images

Contenu



Nom du fichier :

On désigne un fichier sur son disque dur. Le cas échéant, un dialogue propose de le recopier dans le répertoire de l'enquête.

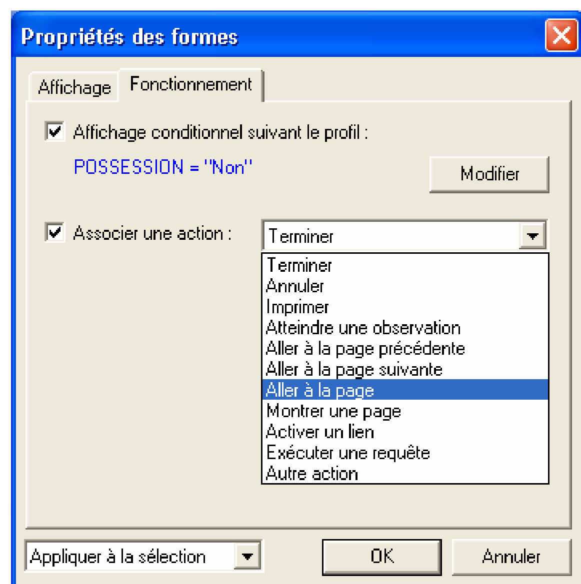
Taille :

Pour adapter l'image originale au formulaire.

Rognage :

Pour montrer une partie seulement de l'image.

Fonctionnement et affichage



Même principe que pour les autres objets.

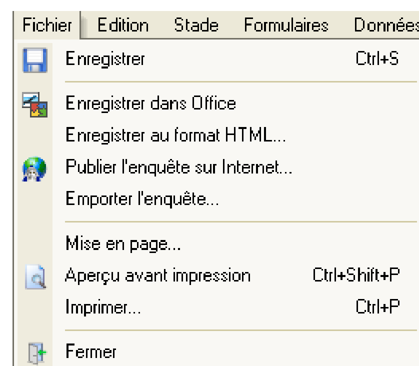
Les illustrations peuvent être utilisées pour la navigation, l'activation d'un lien ou pour mettre fin au questionnaire…

Menu Fichier

Produire le formulaire

Ce menu gère les « sorties » :

- Enregistrement de sauvegarde : **Enregistrer**.
- Production de documents Word ou PowerPoint : **Enregistrer dans Office**.
- Production de pages html : **Enregistrer au format html**.
- Publication de l'enquête sur Internet : **Publier l'enquête sur Internet**.
- Préparation de l'enquête pour le module opérateur ou un pda : **Emporter l'enquête**.
- **Mise en page** : panneau de propriétés de la vue (formulaire ou tableau de bord).
- **Aperçu avant impression** et **Imprimer** : pour vérifier et lancer l'impression de la vue.



Enregistrer dans Office

Transférer le formulaire courant dans l'application choisie (Word – Excel – PowerPoint).



Emporter l'enquête (*)

- Répertoire pour le module opérateur :

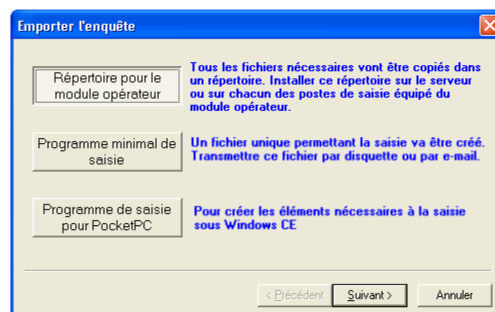
Pour créer les éléments nécessaires à la saisie sur les postes des télé-acteurs et fixer les droits de l'opérateur : saisir, modifier, importer.

- Programme minimal de saisie :

Programme autonome que l'on peut envoyer par mail. Les réponses reviennent dans la boîte électronique indiquée en paramètres. Attention : ce format est souvent interdit par les firewalls anti virus.

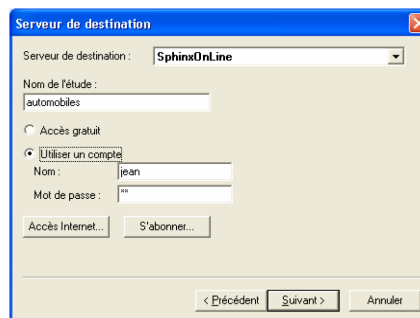
– **Programme de saisie pour PocketPC :**

Créer les éléments nécessaires à la saisie sous Windows CE.

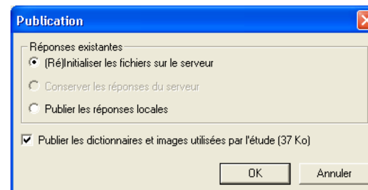


Publier l'enquête sur Internet

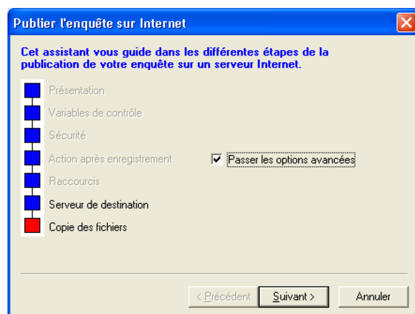
Publication rapide



Un compte de 5 000 unités sur SphinxOnline est ouvert à l'achat de la licence. Il permet de mettre vos questionnaires sur Internet en quelques clics.



Publication gratuite sur le compte public : aucune unité n'est décomptée, mais les données sur ce compte ne sont pas pérennes.



Etude "automobiles"

Vendredi 2 Septembre 2005
SphinxOnLine (public/automobiles)
Ce fichier HTML se trouve dans le répertoire de l'enquête

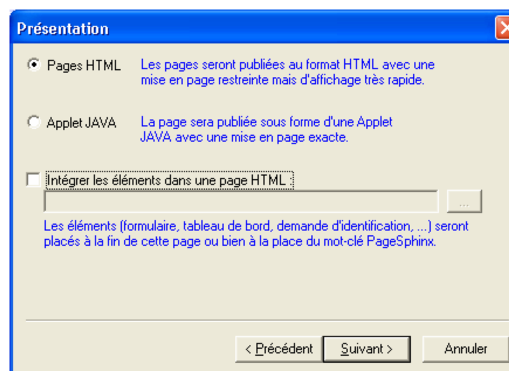
- http://www.sphinxonline.net/public/automobiles/questionnaire.htm
- http://www.sphinxonline.net/public/automobiles/resultats.htm
- http://www.sphinxonline.net/public/automobiles/tables.htm

Lorsque les fichiers ont été publiés, les liens vers les formulaires, les tableaux de bord ou la table des données sont communiqués.

Assistant de publication (*)

Présentation

L'option Html doit être privilégiée lorsqu'on ne connaît pas le type d'accès des correspondants (débit et type de navigateur).



(*) Variables de Contrôle

Clé : Une variable clé est automatiquement générée. Son contenu est communiqué au répondant après enregistrement. Il pourra ainsi revenir sur ses réponses ou continuer la saisie.

Date de saisie, Adresse IP : La date et l'heure de saisie sont enregistrées ainsi que la référence de la machine qui a répondu.

(*) Sécurité :

Pour empêcher les requêtes non autorisées.

(*) Identificateurs de saisie

Les identifiants de saisie sont communiqués à chaque répondant. Ils leur permettent d'accéder au questionnaire.

(*) Action après l'enregistrement :

Préciser le contenu du bouton destiné à mettre fin à la saisie et à lancer l'enregistrement des réponses.

Définir ce qui advient après cette action (rappel des réponses, proposition d'un lien vers la page d'un site ou par exemple, vers les résultats de l'enquête).

(*) Raccourcis :

Ce dialogue définit les liens donnant accès au questionnaire ou aux résultats.

Ce sont des requêtes au serveur d'enquêtes :

Saisir : créer un nouvel enregistrement.

Consulter : voir le contenu des enregistrements existants.

Modifier : accéder à un enregistrement existant et le modifier, une clé est demandée.

Voir les résultats : afficher un tableau de bord.

Voir les tables : lister les données brutes.

La syntaxe des liens est la suivante :

www.sphinxonline.net/compte/enquête/raccourcis.htm

(*) Serveur de destination

SphinxOnline est le serveur par défaut.

La publication peut également se faire sur un autre serveur équipé du logiciel serveur d'enquêtes.

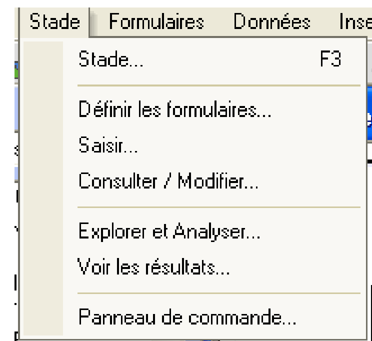
Après avoir choisi le serveur, on précise le nom utilisé pour publier l'enquête, et le compte sous lequel les fichiers seront enregistrés. Ces informations seront reprises dans les liens vers les questionnaires ou les résultats :

www.serveur/compte/nom_etude/ /…

(*) Copie des fichiers

En retour, les liens construits à l'étape **Raccourcis** sont envoyés.

Menu Stade



Définir les formulaires :

Donne accès à l'environnement de travail multimédia.

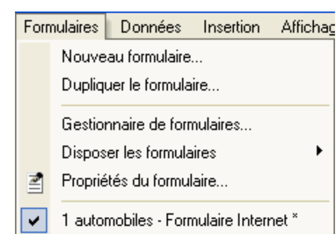
Saisir… :

Pour tester le fonctionnement de la saisie. Renvoie l'interface du module opérateur.

Consulter/ Modifier :

Pour voir l'ensemble des données au travers du formulaire courant.

Menu Formulaire

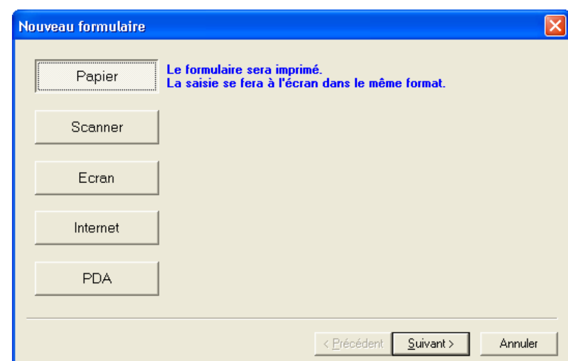


Pour créer un nouveau formulaire, dupliquer un formulaire existant, supprimer ou modifier l'ordre des formulaires dans la liste de la fenêtre formulaires (**Gestionnaire de formulaires**).

Propriétés du formulaire :

Pour accéder au panneau de propriétés du formulaire (format, fonctionnement…).

Nouveau formulaire



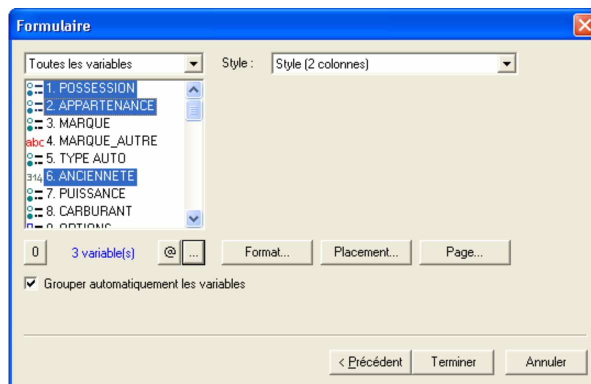
Les formulaires sont répertoriés par média (papier, scanner, …).

Sélection des questions / variables

A la création d'un nouveau formulaire, la totalité des variables est sélectionnée. On obtient un formulaire vierge en cochant la case **0** sous la liste des variables, ce qui permet de construire un questionnaire pas-à-pas.

Style pour choisir un style prédéfini.

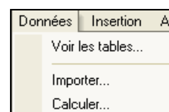
Format, Placement, Page pour adapter le style courant en modifiant les propriétés des questions et les règles de mise en page.



Menu Données

Donne accès au fichier de données brutes pour :

- Le consulter : **Voir les tables**.
- Ajouter de nouvelles observations : **Importer**.
- Créer de nouvelles variables : **Calculer**.



Voir les tables, Importer

Voir les tables :

Pour consulter et / ou modifier les données dans un format tableur ou fiche.

Sur l'ensemble des données, sur une strate ou sur un profil.

Importer :

Accéder à un compte SphinxOnline ou à un serveur d'enquêtes et télécharger les réponses collectées par Internet. Celles-ci viennent s'ajouter aux données courantes ou les remplacer.

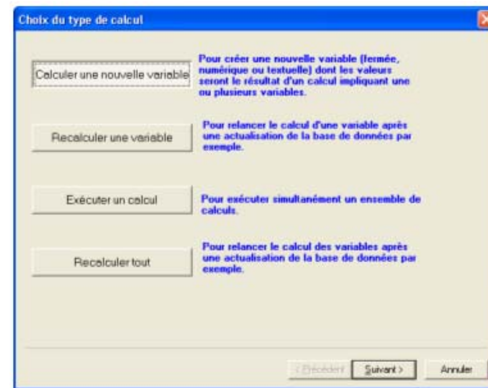
Strates

Variables | Observations | Fiches

Toutes les variables — Noms des modalités / Numéros des modalités

N°	1. POSS...	2. APPA...	3. MAR...	4. MAR...			S.
1	Oui	A quelqu...	Citroën		Occasion	5	4 à 5 CV
2	Oui	A quelqu...	Citroën		Occasion		- de 4 C\
3	Oui	A quelqu...	Autre	Ferrari	Neuve	2	
4	Oui	A vous	Renault		Neuve	4	6 à 7 CV
5	Oui	A vous	Peugeot		Occasion	8	4 à 5 CV
6	Oui	A vous	Renault		Neuve	2	4 à 5 CV
7	Oui	A quelqu...	Renault		Neuve	2	4 à 5 CV
8	Oui	A vous	Peugeot		Occasion	3	4 à 5 CV
9	Oui	A vous	Ford		Neuve	3	4 à 5 CV

Calculer



Pour créer une nouvelle variable par formule ou par programme.

Calculer une nouvelle variable

Avec la calculette ou l'éditeur assisté.

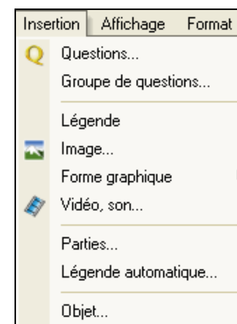
Exécuter un calcul :

Permet de changer à la volée la valeur d'une variable ou de plusieurs variables. Par exemple :

V1=V1/6,65 : convertir des euros en francs

V2=V2*100 : exprimer un ratio en %.

Menu Insertion



Pour insérer différents objets dans un formulaire :

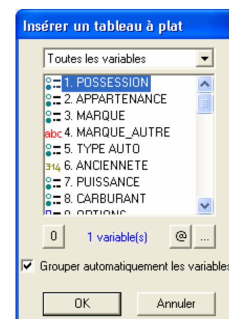
Groupe de questions : Les groupes sont définis au stade de conception du questionnaire.

Légende : Dans une légende, des formules peuvent être ajoutées au texte selon la syntaxe :

[Nom de variable] ou [Fonction] renvoie la valeur de la variable ou la valeur calculée par la fonction.

Par exemple : [Note] / [Moyenne.Note] affiche la note attribuée par le répondant et la valeur moyenne donnée par l'ensemble des répondants.

Question, Groupe de questions, Partie



Questions et groupes de questions

Sélectionner une ou plusieurs questions dans la liste des variables ou dans la liste des groupes définies dans l'élaboration du questionnaire.

Parties

Choisir dans la liste des parties définies dans l'élaboration du questionnaire.

Légendes, images, formes, objets

Légendes :

Une zone s'ouvre dans le formulaire pour écrire la légende.

Formes :

Choisir dans la liste.

Images, Objets :

Ouvrir le fichier correspondant.



Légendes automatiques

Les légendes automatiques permettent d'afficher les caractéristiques du questionnaire définies par les fonctions dont voici quelques exemples :

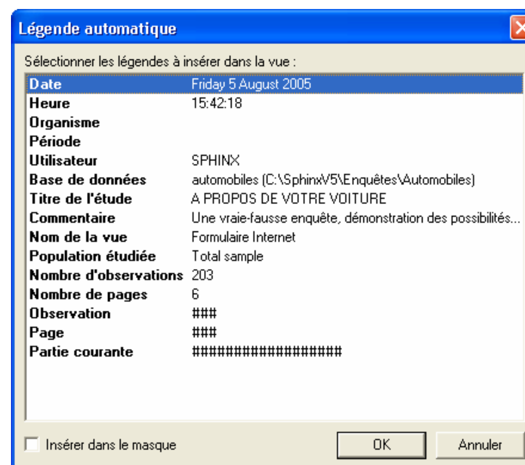
[Nombre d'observations]

[Nombre de pages]

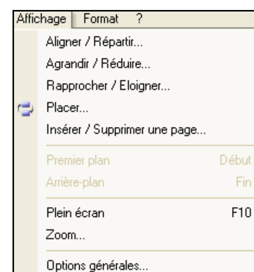
[Observation]

[Page]

[Partie courante]

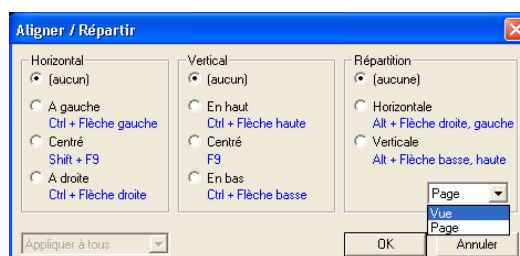


Menu Affichage



Ce menu fixe les règles de placement des objets dans la vue (**Aligner/ Répartir, Agrandir / Réduire, Rapprocher / Eloigner**) ainsi que le découpage des vues en pages (**Insérer / Supprimer**).

Aligner / Répartir, Agrandir / Réduire

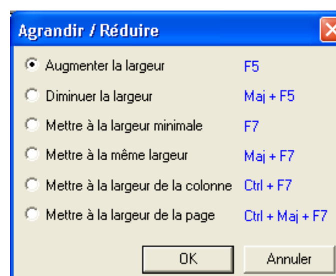


Aligner /Répartir :

Attention :

La disposition des questions dans les formulaires publiés au format html dépend de leur alignement exact.

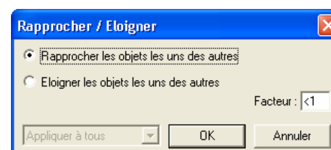
Les raccourcis **Ctrl+flèche** sont le meilleur moyen pour y parvenir.



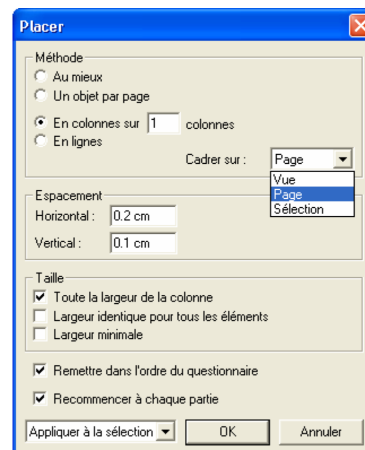
Agrandir / Réduire :

On obtient facilement des formulaires bien présentés en mettant toutes les questions à la même largeur (**Maj+F7**) ou à la largeur de la colonne (**Ctrl+F7**).

Voir les raccourcis dans le menu **? / Raccourcis et Accélérateurs**.



Placer

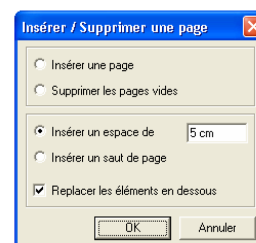


Idéal pour obtenir rapidement une présentation homogène.

Les espacements sont relatifs à l'espace entre deux questions (espace vertical et horizontal) et entre colonnes (espace horizontal).

Les paramètres de taille s'appliquent également à tous les types d'objets.

Insérer / Supprimer



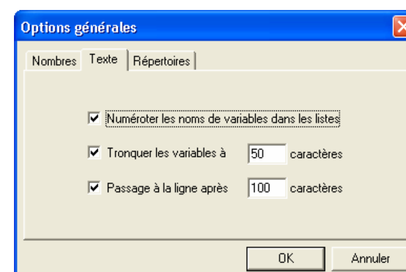
Insérer une page :

Insère une page vide avant la page contenant un objet sélectionné.

Insérer un saut de page :

Fait passer à la page suivante après l'objet sélectionné.

Options générales



Nombres :

Pour fixer le nombre de décimales dans les nombres et pourcentages.

Texte :

Pour fixer les règles d'affichage des variables, de la numérotation, et de découpage des textes.

Menu Format



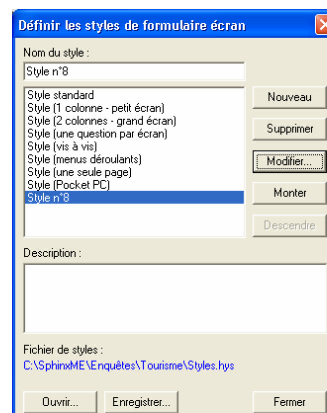
La notion de format renvoie à un ensemble de paramètres définissant les objets d'un formulaire.

Copier le format : Pour copier les paramètres définissant l'objet sélectionné.

Coller le format : Donner à un autre objet, après l'avoir sélectionné, les propriétés de l'objet dont on vient de copier le format.

Choisir comme format : Pour appliquer par défaut les paramètres de l'objet sélectionné aux nouveaux objets insérés dans la vue.

Modifier le format : Pour modifier les paramètres de l'objet sélectionné.



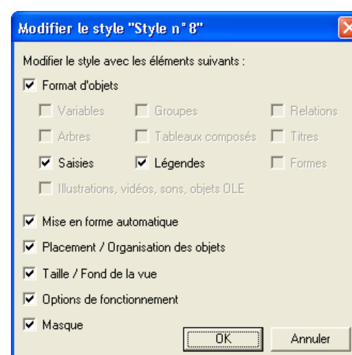
Un style correspond à un ensemble de formats pour le formulaire et ses objets. A la création d'un nouveau formulaire, les questions auront les propriétés du style par défaut ou du style courant.

Appliquer un style :

Donner au formulaire les propriétés du style sélectionné.

Définir les styles :

Pour enregistrer un nouveau style correspondant aux propriétés des objets du formulaire actif. Attention : vérifier que le formulaire de référence ne contient pas 2 objets du même type de format différent.



Icônes et raccourcis pour les formulaires

Gérer l'environnement et échanger

1/ Fermer et revenir au panneau central. 2/ Enregistrer. 3/ Publier sur Internet. 4/ Envoyer dans Office ou OpenOffice (Word, PowerPoint, Excel. 5/ Aperçu de la vue. 6/ Cacher / Afficher la liste des variables. 7/ Zoom sur la vue.	1 2 3 4 5 6 7

Construire le questionnaire

8/ Sélectionner une zone avec la souris. 9/ Insérer la ou les questions sélectionnées. 10/ Insérer une légende, un commentaire. 11/ Insérer une illustration. 12/ Insérer un fichier vidéo ou son. 13/ Supprimer la ou les questions sélectionnées. 14/ Placer automatiquement les questions dans le formulaire (disposition, taille, espacement). 15 / Modifier les propriétés de la vue. 16/ Modifier les propriétés des questions sélectionnées.	8 9 10 11 12 13 14 15 16

13. Tableaux de bord

L'environnement de travail

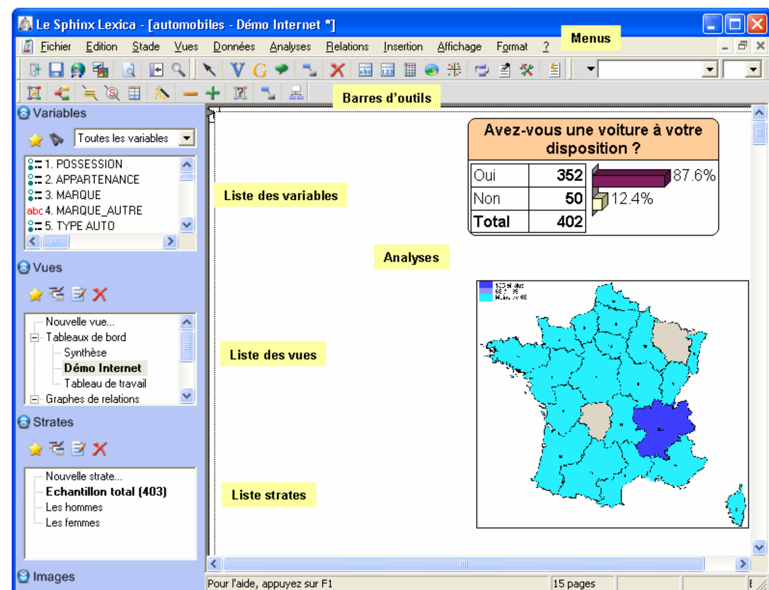
L'environnement de travail multimédia permet de :

- Placer à partir d'une sélection dans la liste des variables, les éléments d'un Tableau de Bord (Tableaux, graphiques, légendes, cartes, graphes arbres…) correspondant à tous types de traitement statistique.

- Créer plusieurs Tableaux de Bord, répertoriés dans la liste des **Vues**.

- Agrémenter les tableaux de bord de légendes et d'illustrations.

Les objets (variables, groupes, graphes) sont placés dans la vue (Tableau de Bord) en utilisant les menus et raccourcis de la barre d'outils.



Propriétés de la vue

Taille



Format :

Choix de la taille de la vue selon le média envisagé.

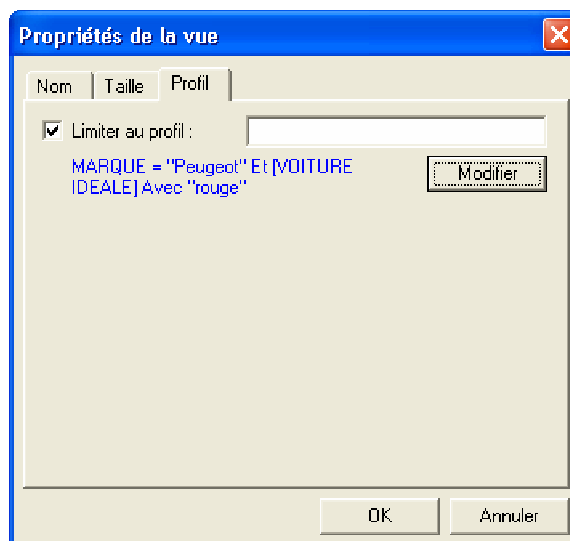
Barre de navigation :

Pour mettre sur les tableaux de bord diffusés par Internet ou par module opérateur, les boutons de passage d'un écran à l'autre.

Activer le lissage :

Le lissage a pour effet de rendre les couleurs transparentes et de permettre la lecture des éléments superposés. Utile notamment pour rendre les cartes factorielles plus lisibles.

Profil



Profil

Pour mettre un filtre sur une vue.

Les éléments du Tableau de Bord sont alors calculés sur les seules observations répondant aux conditions du profil.

Ce profil vient s'ajouter aux filtres des strates et conditionne les profils donnés aux objets (Voir Partie **Propriétés des variables** / **Calculs : limiter au profil**).

Strates

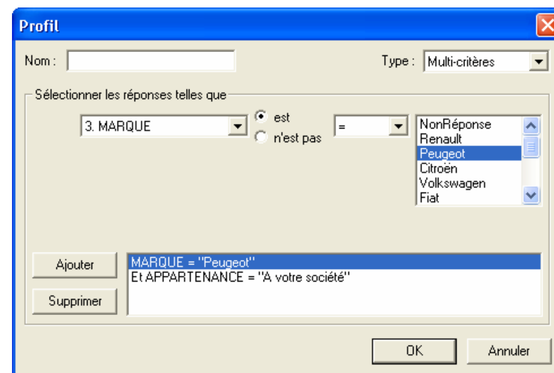


Strate

Une strate est un sous-ensemble de l'échantillon défini par les observations répondant à un profil donné.

Profil :

Dialogue permettant de fixer les conditions auxquelles doivent répondre les observations de la strate.



Changer de strate

Cliquer dans la liste des strates sur la strate à analyser. Les éléments du tableau de bord se trouvent automatiquement calculés sur cette strate.

Attention : les strates définies dans l'environnement classique ne sont pas automatiquement reprises et doivent être définies à nouveau dans l'environnement multimédia **(Nouvelle strate).**

Propriétés des variables

Contenu

La variable est décrite par les éléments sélectionnés dans l'onglet **Contenu** :

Titre de question :

Le libellé et le numéro de la question et le nom de variable.

Description :

Une zone de texte présentant un commentaire.

Valeurs remarquables :

Indicateurs statistiques résumant la variable.

Tableau à plat :

Tableau d'effectifs et de pourcentages. La **largeur des colonnes** du tableau est fixée en cm. Le texte s'inscrit sur plusieurs lignes s'il dépasse cette largeur. Ceci permet de rendre homogène la largeur des tableaux.

Pour déterminer une largeur différente à chaque colonne du tableau, on indique successivement les largeurs respectives de chaque cellule, en les séparant par des points virgules.

Graphique :

Graphe sélectionné dans la liste proposée.

Le texte des libellés, variables et modalités peut être modifié avec le bouton **...** présent à droite du nom de la variable. Ces modifications sont effectuées dans le dialogue suivant.

Valeurs

La variable est décrite par des indicateurs présentés dans la zone des valeurs au-dessus du tableau et sélectionnés dans l'onglet **Valeurs** :

- Taux de réponse : réponses effectives.

- Mode : valeur la plus fréquente.

- Nombre de valeurs différentes.

- **Effectif moyen.**

- Valeurs numériques : moyenne, écart-type, médiane, minimum - maximum, somme, percentiles.

- Moyenne corrigée : moyenne calculée pour les seules valeurs comprises dans l'intervalle défini par les percentiles sélectionnés.

- **Commentaire** : pour afficher si nécessaire un commentaire automatique relatif à certains calculs.

Calcul : profils et non-réponses

- **Limiter au profil** :

Donne les valeurs du tableau pour les seules observations correspondant au profil.

- **Ignorer les non-réponses** :

Les pourcentages sont calculés sans tenir compte des non-réponses.

- **Profil multi-critères :**

Sélectionner **Multi-critères** dans le menu déroulant **Type**.

Pour ajouter plusieurs conditions au profil en choisissant comment les combiner (opérateurs et – ou – sauf).

Calcul : variables nominales (questions fermées, échelles)

Variables fermées multiples

Le calcul peut être limité à chacun des rangs d'une variable à réponses ordonnées : **Uniquement la réponse au rang…**

Variables fermées échelles

Quand l'option **Traiter les échelles en nombres** est cochée, la variable est considérée comme une numérique. Les valeurs associées à chaque échelon sont modifiables en cliquant sur le bouton **…**

Calcul : Variables numériques

Mettre les nombres en classes :

Si la case n'est pas cochée, aucun tableau ne s'affiche. Quand l'option **Mettre les nombres en classes** est cochée, le tableau est construit selon le nombre de classes indiqué et sur des classes **de même valeur** ou **de même amplitude**.

Autour de la moyenne propose un découpage dont l'amplitude se fixe en écart-type.

Personnalisées permet de définir les classes en fixant directement les seuils.

Calcul : variables dates et codes

Variables codes :

Pour interpréter un code :

-Indiquer le Numéro des caractères à prendre en considération : par exemple les 2 premiers pour le code postal.

-Associer le fichier du dictionnaire à utiliser pour regrouper automatiquement les codes en catégories.

Variables dates :

Les dates peuvent être regroupées en intervalles années, mois… ou en périodes personnalisées, dont on indique les dates charnières en les séparant par des points virgules.

Calcul : variables textes

Mise en classes des réponses :

Regroupe les réponses identiques et en donne le nombre.

Rappel : si la réponse contient plusieurs lignes ou des points virgules, elle sera éclatée en autant de lignes.

Mise en classes des mots :

Donne, pour chaque mot, le nombre de fois où il apparaît dans les réponses. Le bouton **...** permet d'ignorer les mots outils (dictionnaire des mots outils). On peut également ne faire apparaître que certains mots entrés dans une liste ou présents dans un dictionnaire.

Nombre de mots :

Renvoie un nombre total ou moyen de mots selon les indicateurs numériques choisis dans l'onglet **Valeurs**.

Calculs : simplifier le tableau

Pour limiter l'affichage des lignes d'un tableau à :

- Un **effectif minimum** ou un **pourcentage minimum** : les modalités présentant des effectifs ou pourcentages inférieurs aux chiffres indiqués n'apparaissent pas dans le tableau.

-Un **nombre maximum de lignes**.

Ou pour simplifier le tableau en regroupant ou en supprimant des lignes : **Simplification personnalisée.**

La syntaxe est la suivante :

Françaises = Renault = Peugeot = Citroën	pour regrouper les trois marques françaises.
(Volvo ; Toyota)	pour supprimer Volvo et Toyota.
10-20	pour n'afficher que les lignes 10 à 20.

- Nom du groupement=Nom ou numéro des modalités à regrouper.
- Suppression : noms des modalités entre parenthèses.
- Les éléments sont séparés par des ' ;'

Simplifier
Pourcentage minimum
0%
Regrouper sous
autres
Supprimer
Regrouper sous

Simplifier
Simplification personnalisée
Pas important=Pas du tout=Plutôt non;Moyen=3;Import
...

Simplification personnalisée
Pas important=Pas du tout=Plutôt non
Moyen=3
Important=4=5

Tableau

Pour fixer la présentation du tableau : totaux marginaux, mode de tri des lignes, possibilité de limiter à un certain nombre de lignes.

Contenu des cellules :

Pour indiquer les éléments affichés dans le tableau : **Effectif** et / ou **Pourcentage** et / ou **Intervalle de confiance**.

Pour déterminer le mode de calcul des pourcentages : taux calculés **sur les observations** ou **sur les citations**.

Attention, pour les questions à réponses fermées multiples, le pourcentage doit être par défaut calculé par rapport au nombre d'observations.

Référence : pour comparer les effectifs ou pourcentages du tableau :

-Aux résultats sur l'échantillon total (**Toutes les observations**).

-Aux **observations de la vue**.

-A des **données externes**.

On utilise les deux premières possibilités quand le calcul du tableau est effectué sur un profil particulier.

Tableau : mise en forme conditionnelle

Pour souligner dans un tableau les chiffres ou modalités remarquables et orienter la lecture en fonction de conditions énumérées dans la liste **Mises en forme** :

-Les éléments ayant la plus grande/petite valeur.

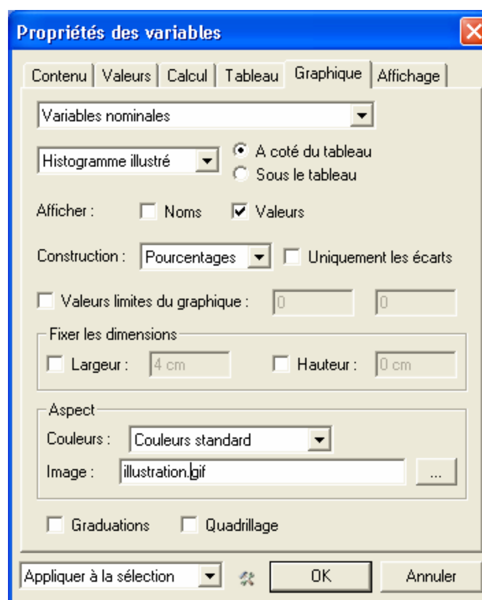
-Les éléments sur/sous représentés.

-Les éléments compris dans un intervalle.

-Les éléments répondant à un certain profil.

Le format (couleur, police) indique comment faire ressortir les cellules, lignes ou colonnes.

Graphique



Choix du graphique :

Secteur, Histogramme, Barre, Courbe, Aire, Anneau, Radar, Boite à moustache (numérique uniquement), Histogramme illustré.

Afficher :

Pour présenter les **noms** des modalités et les **valeurs** associées sur le graphique.

Construction :

Uniquement sur les écarts pour représenter l'écart à une référence (onglet tableau).

Echelle du graphique :

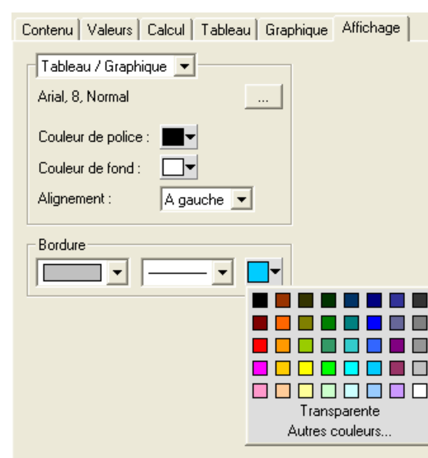
Fixer les **valeurs limites** et la taille des graphiques (**Largeur et Hauteur**) pour harmoniser la représentation.

Aspect :

Image : pour les histogrammes illustrés, associer le ou les fichiers images associés à chaque modalité ('**...**').

Couleurs : pour choisir le bon mode d'expression.

Affichage



Les polices, couleurs, alignements sont définis après avoir choisi dans la liste déroulante la zone concernée **Nom** – **Description** - **Tableau/Graphique** - **Valeurs**.

Propriétés des groupes

Contenu

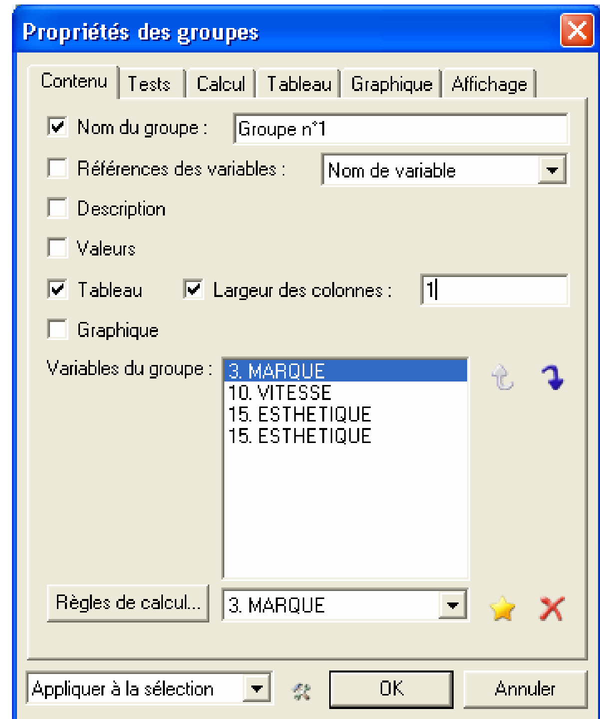
Variables et types de tableaux :

Pour ajouter une variable dans le groupe, cliquer sur l'étoile en bas à droite de la liste, et fixer son mode de calcul (**Règle de calcul**). Par défaut le mode est celui du style. La nature du tableau de groupe dépend de ces choix :

- Tableau de valeurs moyennes si le tableau ne contient que des numériques (échelles en nombre et numérique non mises en classes).

- Tableau croisé d'effectifs si uniquement des nominales (échelles non numériques et numériques en classes).

- Tableau de moyennes croisées si variables nominales et numériques.



Ordre des variables et présentation du tableau :

Les flèches permettent de changer l'ordre des variables dans les tableaux. La variable en tête est le pivot des tableaux croisés. Pour les tableaux de moyennes, choisir le placement des numériques (en ligne ou colonne) dans l'onglet **Tableau**. Le groupe est alors décrit par les éléments sélectionnés dans l'onglet **Contenu** :

- Nom du groupe : la description de l'analyse.

- Référence des variables : l'affichage du nom ou du libellé des variables analysées.

- Description : une zone de commentaire.

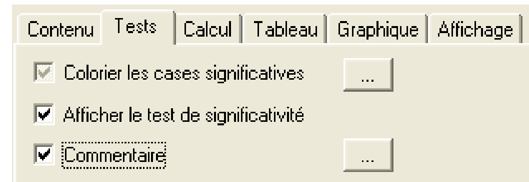
- Tableau : le tableau de groupe tel que défini ci-dessus, et présenté dans une largeur de colonne que vous déterminez en cm (les libellés sont alors automatiquement ajustés sur une ou plusieurs lignes).

- Graphique : graphe sélectionné dans la liste.

Tests

Pour afficher le résultat des tests pertinents selon la nature des variables groupées :

- Test du Chi2 pour un tableau croisé ;
- Analyse de variance pour un tableau de moyennes croisées ;
- Corrélation pour le croisement de deux variables numériques.
- Pour représenter le résultat des tests statistiques, vous pouvez :
- **Colorier les cases significatives** : pour mettre en évidence les résultats significatifs.
- **Afficher le test de significativité** : pour indiquer le test statistique relatif à ce groupe.
- **Afficher le commentaire** : pour afficher d'autres informations prises en considération dans le test statistique.



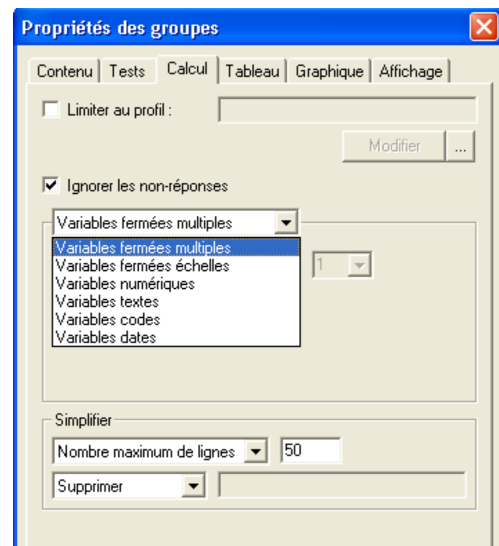
Calcul

Identique à l'onglet **Calcul** du panneau des **Propriétés des variables**.

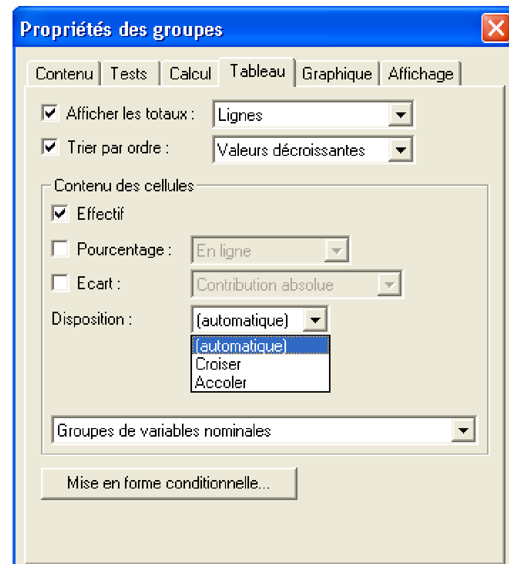
Il permet de spécifier la manière dont chaque variable du groupe intervient dans les calculs en fonction de son type.

Pour différencier les règles de calcul de 2 variables d'un même type, utiliser le bouton **Règles de calcul** de l'onglet **Contenu**.

La restriction à un profil, la gestion des non-réponses et les règles de simplification sont identiques à celles décrites dans la partie **Propriétés des variables**.



Tableau



Présentation

Afficher les totaux :

Les tableaux de groupes intègrent une ligne et / ou une colonne de total.

Trier par ordre :

Les tris s'opèrent par rapport aux valeurs effectifs ou pourcentages de la première colonne du tableau.

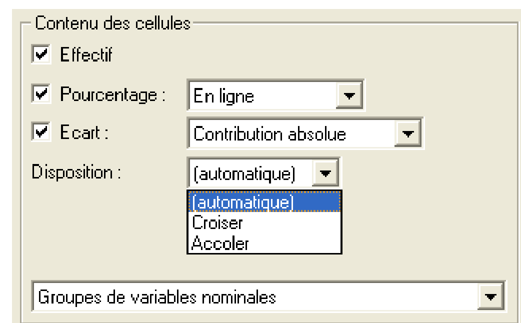
Contenu des cellules :

Il dépend de la nature du tableau (groupe de numériques, groupe de nominales, groupe de numériques et nominales).

Mises en forme conditionnelle :

Identique aux options présentes dans l'onglet **Tableau** des **Propriétés des variables**.

Tableau : Contenu des cellules pour un groupe de nominales



Pourcentage :

Pourcentages calculés par rapport au total des lignes, des colonnes ou de tout le tableau.

Pourcentages calculés par rapport au nombre d'observations des modalités en lignes ou en colonnes.

Ecart :

Pour afficher les éléments du calcul du Chi2, effectifs théoriques, écarts, contributions.

Disposition :

Croiser : la première variable (selon l'ordre indiqué dans l'onglet **Contenu**) avec les autres. Ses modalités figurent en colonne.

Accoler : tris à plat présentés en ligne dans un même tableau. N'a de sens que si les modalités sont identiques (batterie d'échelles).

Tableau : Contenu des cellules pour un groupe de numériques

Choisir les indicateurs à faire figurer dans le tableau : moyenne, écart-type, minimum - maximum, somme, pourcentage, médiane, effectif.

Référence

Pour comparer les valeurs du tableau aux valeurs obtenues sur un autre ensemble de données. Dans le cas de plusieurs indicateurs, la référence n'est donnée que pour les valeurs de la première colonne.

Contenu des cellules
Moyenne
Ecart-type
Min - Max
Somme
Pourcentage
Médiane
Effectif
Référence :
Toutes les observations
Toutes les observations
Observations de la vue
Profil d'observations
Groupes de variables numériques

Tableau : Contenu des cellules pour un groupe de numériques et nominales

Choisir les indicateurs à faire figurer dans le tableau : moyenne, écart-type, somme, part, effectif, t de Student.

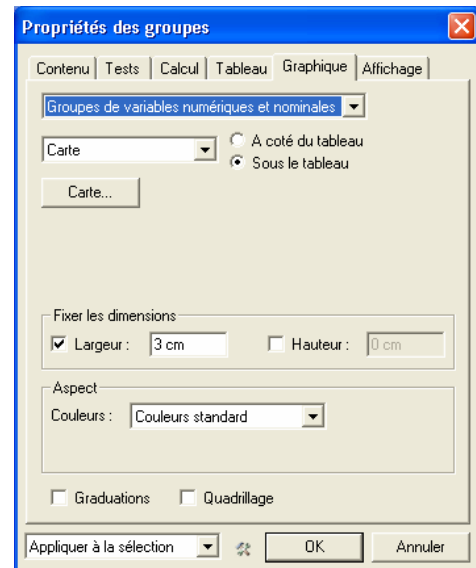
Placement des variables numériques :

Indiquer la disposition des variables numériques : en ligne ou en colonne.

Contenu des cellules
Moyenne
Ecart-type
Somme
Part
Effectif
t de student
Placement des variables numériques :
(automatique)
(automatique)
En ligne
En colonne
Groupes de variables numériques et nominales

Graphiques

Choix des graphiques :



Histogrammes, barres, baromètres, courbes, aires, radars, boites à moustache, secteurs, cartes.

Comme pour les propriétés de variables, on fixe les éléments du graphique, ses échelles, ses dimensions et son aspect.

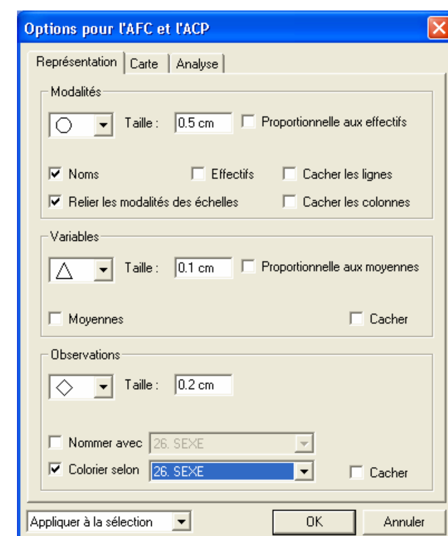
Cartes factorielles :

La représentation des tableaux sous forme de cartes fait l'objet d'un ensemble d'onglets spécifiques.

Ces cartes permettent de représenter dans un plan les valeurs de 2 variables ou plus en utilisant les techniques d'analyse factorielle.

L'icône de la barre d'outils permet d'afficher directement ces cartes.

Cartes factorielles : Représentation des éléments (*)



Onglet **Représentation** du panneau **Options** pour l'AFC :

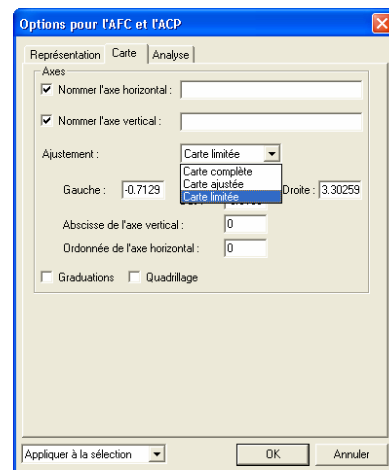
Pour régler l'affichage des éléments de la carte : **Modalités** (AFC), **Variables** (ACP) et **Observations** (AFC ou ACP).

Certains réglages se font ailleurs :

Pour rendre la carte plus lisible, décocher la case **Largeur** dans l'option **Fixer les dimensions** de l'onglet **Graphique** (**Panneau de propriétés des groupes**). On peut ainsi directement adapter la taille de l'objet.

Pour rendre les superpositions transparentes, cliquer **Activer le lissage** dans l'onglet **Taille** du panneau de **Propriétés de la vue**.

Carte factorielle : cadrage de la carte (*)



Onglet **Carte :**

Nommer l'axe horizontal et **Nommer l'axe vertical :**

Pour donner une dénomination aux deux axes.

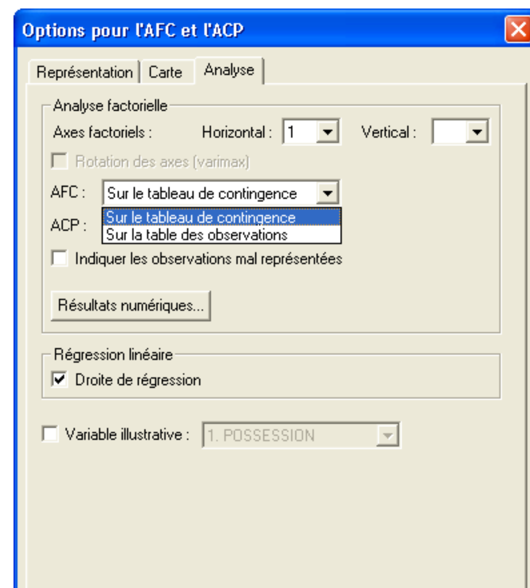
Ajuster ou réduire la carte

Ajuster : sur les points extrêmes.

Limiter : pour ne voir qu'une partie de la carte : définir les coordonnées limites à gauche et à droite, et le centre de la carte (**Abscisse de l'axe vertical** et **Ordonnée de l'axe horizontal**).

Le fond de la carte peut être agrémenté de **Quadrillage** et **Graduations**.

Cartes factorielles : Type d'analyse (*)



Axes factoriels :

Sélectionner les axes pour obtenir la carte correspondante.

Choix de la méthode :

Sélectionner **Sur la table des observations** pour analyser le tableau individus / modalités (AFCM) et **Sur le tableau de contingence** pour analyser le tableau croisé.

Variables illustratives :

Dans le cas d'une analyse menée sur la table des observations. Ces variables n'interviennent pas dans le calcul des axes mais permettent de situer les modalités de la variable correspondante.

Résultats numériques :

Affiche les paramètres de l'analyse :

Tableau de Burt ou matrice de corrélation.

Variances expliquées par les axes.

Contribution, cosinus2, coordonnées des points.

Propriétés des tableaux composés (*)

Tableau construit sur des variables (*)

Un tableau composé est un tableau dont les lignes, les colonnes et le contenu des cellules sont spécifiés par l'utilisateur dans l'onglet **Tableau**.

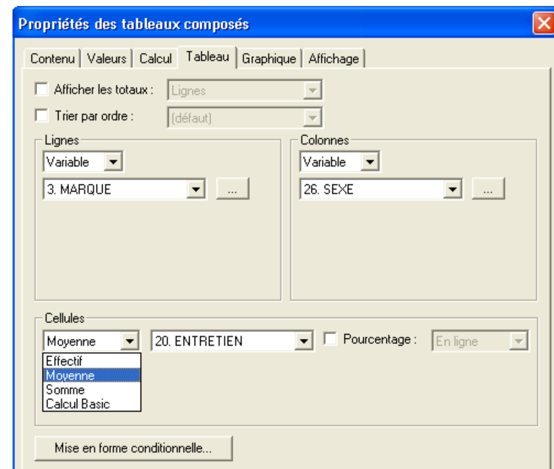


Tableau construit sur des variables

Lignes : choisir la variable dont les modalités définiront les lignes du tableau.

Colonnes : choisir la variable dont les modalités définiront les colonnes du tableau.

Cellules :

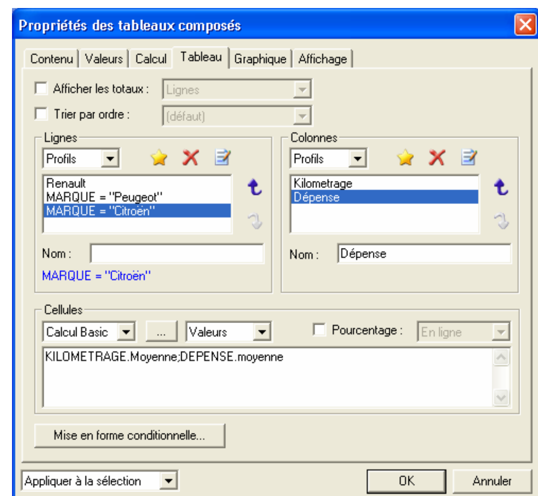
-**Effectif** : donne le tableau croisé des 2 variables.

-**Moyenne** ou **Somme** : conduit à choisir une troisième variable numérique dont la valeur sera calculée pour chaque cellule.

-**Calcul Basic** permet de définir un calcul qui sera effectué pour chaque cellule du tableau.

Tableau construit par profil (*)

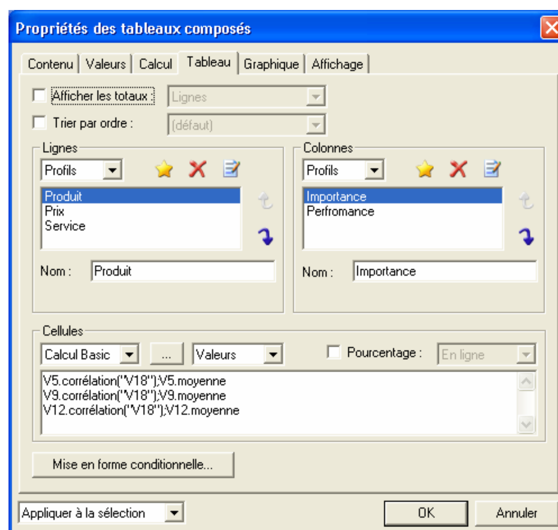
Lignes : choisir **Profils**, écrire la condition définissant la première ligne et lui donner un nom et procéder de la même manière pour les différentes lignes du tableau.



Colonnes : choisir **Profils** et **Tous** si on ne veut pas appliquer d'autre restriction. Nommer la première colonne et procéder de la même manière pour toutes les colonnes du tableau.

Cellules : dans l'exemple ci-contre, on choisit **Calcul Basic** pour entrer ensuite les calculs à effectuer sur chaque ligne. Ici, en colonnes, figurent le kilométrage et la dépense moyenne pour les différents profils d'individus définis en ligne.

Tableau construit par énumération des cellules (*)



L'exemple ci-contre correspond à la construction d'un tableau Importance / Performance.

Lignes : Les calculs de chaque ligne seront effectués pour tous les individus (profil **Tous**). Chaque ligne fait l'objet de calculs mettant en jeu le critère qui lui donne son nom.

Colonnes : On crée deux colonnes intitulées Importance et Performance.

Cellules :

Les calculs effectués pour chaque ligne sont énumérés cellule par cellule :

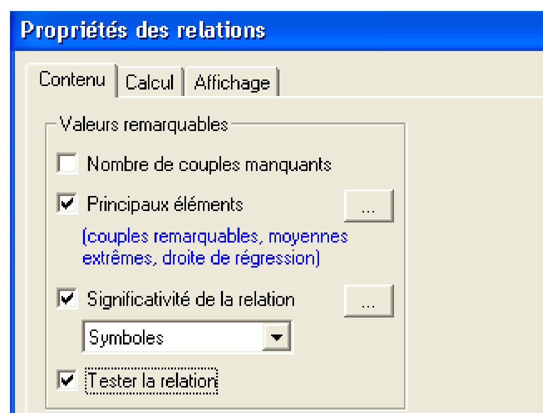
L'importance : Corrélation de V5 (Produit) avec V18 (la satisfaction globale).

La performance : valeur moyenne de V5.

… et ainsi de suite pour chaque ligne.

Propriétés des relations (*)

Contenu (*)



L'onglet **Contenu** permet de documenter les caractéristiques d'une relation entre deux variables : le trait tracé entre les deux variables indique la nature de la relation et son intensité, qui sont également décrits par les indicateurs suivants :

Principaux éléments :

Couples de modalités remarquables ou moyennes extrêmes ou droite de régression. Les règles utilisées pour calculer ces éléments sont modifiables en cliquant sur **...** (Voir Partie Statistiques)

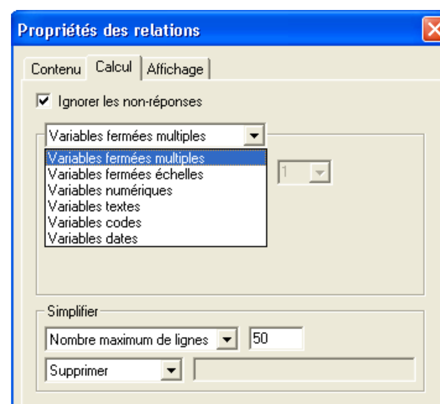
Significativité de la relation :

Résultats du test appliqué compte tenu de la nature des deux variables. Les symboles sont paramétrables en cliquant sur **...** (Voir Partie Statistiques).

Tester la relation :

Pour tracer en pointillés les relations non significatives et relancer le test à chaque affichage.

Calcul (*)



L'onglet **Calcul** permet de définir la manière de traiter les variables selon leur type :

Variables fermées multiples :

Calcul global ou selon le rang.

Variables fermées échelles :

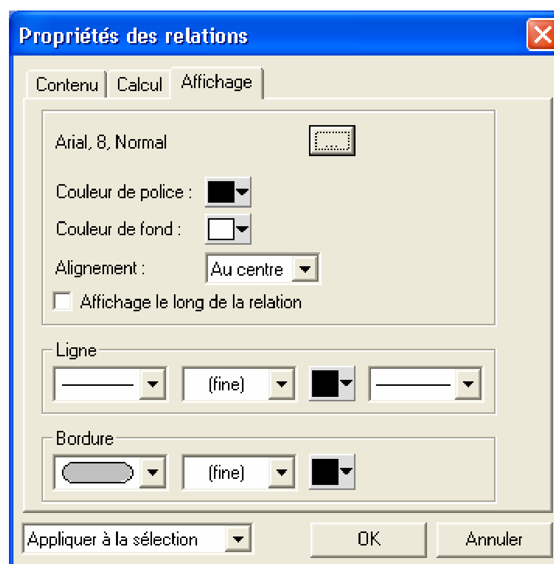
Traitées comme nominales ou comme numériques (échelles en nombres).

Variables numériques :

Traitées comme numériques ou comme nominales (mises en classes).

Les choix effectués à ce niveau déterminent le test effectué et peuvent changer le résultat.

Affichage (*)



Affichage :

Pour donner au trait symbolisant les relations leur apparence, en déterminant la taille et la couleur des polices et le style du trait.

Pour obtenir un trait plus ou moins épais selon la significativité, cocher **Modifier l'épaisseur d'une relation selon sa significativité** dans le dialogue **Mise en forme automatique** (dernière icône à droite dans la barre d'outils).

Statistiques (*)

Les paramètres statistiques sont modifiables à partir de l'onglet **Contenu** du panneau de propriétés des relations en cliquant sur **...** ou dans l'article **Statistique**s du menu **Relations**.

Seuils :

Définir pour chaque test le seuil à partir duquel la relation est considérée comme très significative (TS), Significative (S), Peu significative (PS), Non significative (NS).

Dans les graphes de relations, seules les relations TS et S apparaissent en traits pleins. Les autres n'apparaissent pas ou sont symbolisées par des pointillés (Option **Tester la relation** présente dans l'onglet **Contenu**).

Calculs stricts :

Les tests ne sont effectués que si les règles de calculs stricts sont vérifiées. Quand ce n'est pas le cas, la relation est marquée du symbole ' X '.

Principaux éléments :

Pour fixer les conditions de sélection des éléments remarquables.

Cases couleurs :

Pour fixer les seuils utilisés pour l'affichage des cases significatives dans les tableaux croisés ou de moyennes.

Seuils | Calculs stricts | Principaux éléments | Cases en couleurs

Symboles de significativité des relations

	TS	S	PS	NS
Nominale / Nominale (chi²)	1%	5%	15%	
Nominale / Numérique (anova)	1%	5%	15%	
Numérique / Numérique (corr.)	0.8	0.6	0.4	

Les seuils sont utilisés pour afficher les symboles de significativité des relations.
Les relations non calculées (cf Calculs stricts) sont marquées 'X'.

Seuils | Calculs stricts | Principaux éléments | Cases en couleurs

Conditions de calcul d'une relation

☑ Appliquer les règles de calculs stricts

Pour les relations Nominale / Numérique et Nominale / Nominale, chacune des modalités doit avoir un effectif minimum de : 30

Pour les relations Nominale / Nominale le pourcentage de couples d'effectif théorique inférieur à 5 doit être inférieur à : 20%

Les relations ne répondant pas à ces critères ne seront pas calculées et seront marquées 'X'.

Seuils | Calculs stricts | Principaux éléments | Cases en couleurs

Affichage des principaux éléments d'une relation

☑ Uniquement pour les relations significatives

Effectif minimum : 10

Nominale / Nominale

☑ Uniquement les couples d'effectif supérieur à l'effectif théorique

Nombre maximum d'éléments affichés : 5

Seuils | Calculs stricts | Principaux éléments | Cases en couleurs

Coloriage des cases significatives dans les tableaux

Nominale / Nominale (chi²)

Les cases coloriées sont celles qui contribuent le plus au chi2 jusqu'à concurrence de : 60%

Nominale / Numérique (anova)

Les cases coloriées ont une moyenne différente de l'ensemble au seuil de : 95%

Les cases ne sont coloriées que si la relation est significative.
Les cases sont bleues si la valeur est supérieure à la valeur théorique, rose sinon.

Arbres de décision : propriétés des nœuds (*)

Contenu (*)

Un arbre de décision est une représentation de la manière dont les données peuvent être décomposées. Un nœud de l'arbre correspond à un profil (ensemble d'individus) qui sera décomposé en sous-ensembles (nouveaux nœuds) définis par de nouveaux profils et ainsi de suite.

Valeurs remarquables :

Pour spécifier les indicateurs caractérisant chaque nœud.

Illustrer avec la variable :

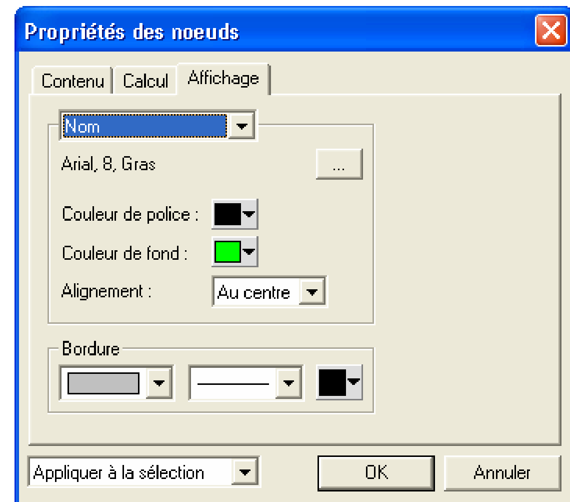
Permet de représenter la variable choisie pour la population de chaque nœud.

Calcul (*)

Pour définir le profil correspondant au nœud.

Affichage (*)

Pour fixer l'aspect de l'arbre, en déterminant le style des polices utilisées, et celui des nœuds.

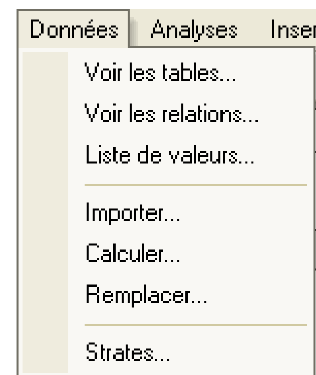


Menu Données

Donne accès à la table des données pour la consulter, la modifier ou l'enrichir.

Pour importer des données, créer de nouvelles variables par calcul ou remplacer les valeurs (modalités ou mots d'une variable).

Pour créer une nouvelle strate.



Voir les tables

Observations :

Affiche la table des données.

Cliquer sur une cellule pour la modifier, sur le titre d'une colonne pour la trier.

Pour restreindre l'affichage à quelques variables, taper leurs numéros en les séparant par des points virgules : 1 ; 3 ; 6-8 par exemple.

Strates

Variables | Observations | Fiches | Liste de valeurs | Résumé

Toutes les variables | Remplacer... | Supprimer | Annuler | des modalités

N°	1. POSSESSI...	2. APPARTE...	3. MARQUE	4. MARQUE_...	5. TYPE AUTO	6. ANCIENNE...	7. PUISSANCE	8. CARBURA...	9. OPTIONS
1	Oui	A quelqu'un d...	Citroën		Occasion	5	4 à 5 CV	Sans plomb	Ordinateur
2	Oui	A quelqu'un d...	Citroën		Occasion		- de 4 CV	Sans plomb	Autres
3	Oui	A quelqu'un d...	Autre	Ferrari	Neuve	2		Gasoil	Airbags
4	Oui	A vous	Renault		Neuve	4	6 à 7 CV	Essence	Climatisation
5	Oui	A vous	Peugeot		Occasion	8	4 à 5 CV	Essence	Autres
6	Oui	A vous	Renault		Neuve	2	4 à 5 CV	Gasoil	Airbags
7	Oui	A quelqu'un d...	Renault		Occasion	2	4 à 5 CV	Gasoil	Airbags
8	Oui	A vous	Peugeot		Occasion	3	4 à 5 CV	Sans plomb	Airbags
9	Oui	A vous	Ford		Neuve	3	4 à 5 CV	Essence	Climatisation
10	Oui	A vous	Mercedes		-	4	8 à 9 CV	Essence	Climatisation
11	Oui	A quelqu'un d...	Renault		Neuve / Occasion	5	4 à 5 CV	Gasoil	Autres
12	Oui	A vous	Volkswagen		Occasion		4 à 5 CV	Gasoil	Jantes alu
13	Oui	A vous	Opel		Occasion	2	4 à 5 CV	Gasoil	Climatisation
14	Oui	A vous	Autre	Seat	Occasion		6 à 7 CV	Essence	Airbags
15	Oui	A votre société	Renault		Neuve		6 à 7 CV	Gasoil	Airbags
16	Oui	A vous	BMW		Neuve	1	8 à 9 CV	Essence	Autres
17	Non								
18	Oui	A vous	Volkswagen		Occasion	3	6 à 7 CV	Sans plomb	Jantes alu
19	Oui	A vous	Ford		Neuve	4	4 à 5 CV	Sans plomb	Autres

Les boutons **Strates** et **Profils** en haut à gauche permettent de limiter l'affichage aux seules observations correspondant au profil indiqué ou à la strate sélectionnée.

Fiches :

Pour voir les données observation par observation.

Voir les relations

Relations :

Permet de visualiser la matrice de toutes les relations établies selon les règles de calcul choisies et les valeurs de seuil fixées pour chacun des tests.

Strates

Relations

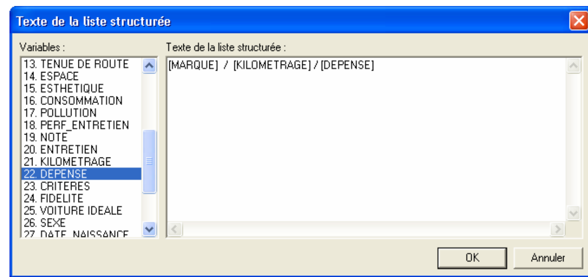
Variable	Nb. Signif.	1. POSSESSION	2. APPARTENANCE	3. MARQUE	5. TYPE AUTO	6. ANCIENNETE	7. PUISSANCE	8. CARBURAN
7. PUISSANCE	22	NS	NS	TS	TS	TS		TS
28. CSP	19	TS	TS	NS	TS	TS	TS	PS
9. OPTIONS	18	TS	NS	PS	TS	TS	TS	S
24. FIDELITE	16	NS	NS	S	TS	TS	S	TS
5. TYPE AUTO	16	NS	PS	PS		TS	TS	TS
3. MARQUE	16	NS	NS		PS	TS	TS	TS
31. ToutesMarques	15	NS	NS	TS	NS	TS	TS	PS
26. SEXE	15	S	TS	NS	NS	NS	S	NS
8. CARBURANT	15	NS	PS	TS	TS	TS	TS	
22. DEPENSE	13	NS	TS	TS	TS	NS	TS	S
13. TENUE DE ROUTE	13	NS	NS	TS	TS	S	TS	S
10. VITESSE	12	NS	NS	TS	S	S	TS	TS
6. ANCIENNETE	12	NS	S	TS	TS		TS	TS
21. KILOMETRAGE	10	NS	TS	PS	TS	NS	TS	S
19. NOTE	10	NS	NS	TS	TS	S	TS	TS
15. ESTHETIQUE	10	NS	PS	TS	TS	PS	TS	S
20. ENTRETIEN	9	NS	TS	S	NS	NS	TS	NS
23. CRITERES	8	S	NS	NS	PS	NS	NS	S
14. ESPACE	8	NS	NS	TS	TS	NS	TS	NS

Règles de calcul... | Statistiques...

Listes de valeurs

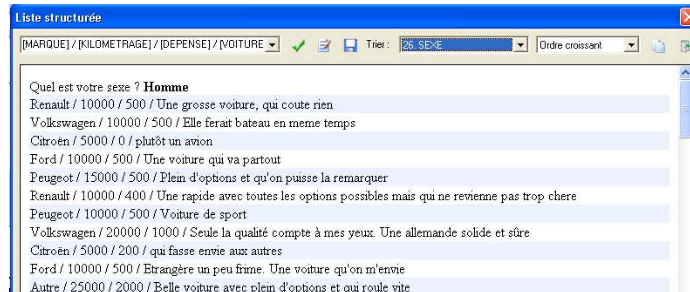
Pour dresser la liste des valeurs pour les variables sélectionnées.

Le format de la liste est éditable en ajoutant du texte entre le nom des variables marqué entre crochets.



Présentation des listes

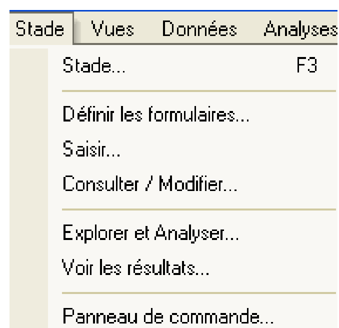
La liste peut éventuellement être triée selon les modalités d'une autre variable et/ou limitée aux individus répondant à un certain profil.



Menu Stade

Explorer et Analyser :

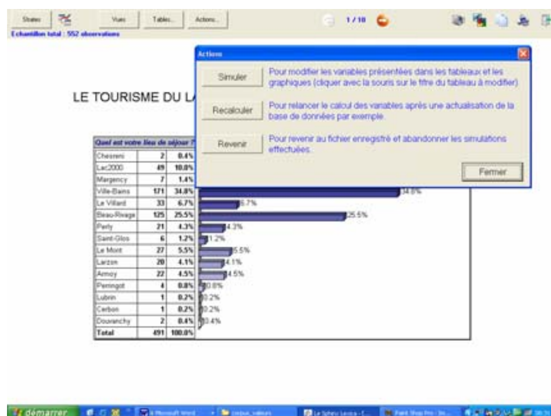
Pour se placer dans l'environnement de construction des tableaux de bord.



Voir les résultats :

Pour consulter les tableaux de bord dans l'interface du module opérateur.

Voir les résultats



Les boutons situés en haut de l'écran permettent de sélectionner les résultats à afficher.

Vues :

Pour sélectionner un tableau de bord.

Strates :

Pour voir les résultats sur l'une des strates, ou définir un nouveau profil.

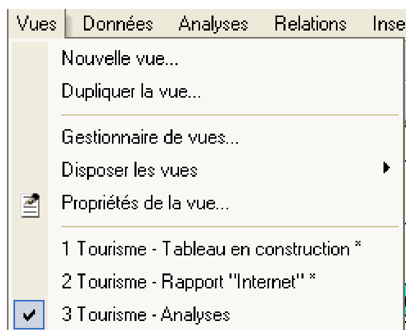
Tables :

Pour afficher les données brutes.

Actions :

Pour changer le contenu d'un tableau ou recalculer les variables.

Menu Vues



Pour :

Créer une nouvelle vue ou dupliquer une vue existante (**Nouvelle vue** et **Dupliquer la vue**).

Organiser la liste des vues et leur affichage avec les articles **Gestionnaire de vues** et **Disposer les vues**.

Définir les propriétés de la vue courante (**Propriétés de la vue**).

Les vues ouvertes apparaissent dans les derniers articles du menu.

Nouvelle vue

Pour choisir le type d'analyse ou de présentation envisagé :

Tableau de bord :

Pour présenter des résultats statistiques sous forme de tableaux et graphiques.

Graphe de relations :

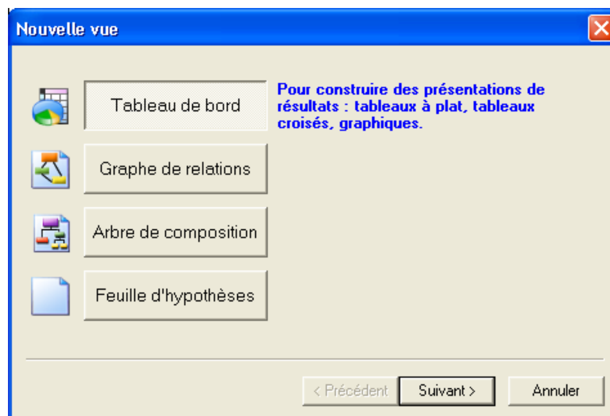
Pour explorer l'existence des relations statistiques entre les variables de l'étude.

Arbre de composition :

Pour construire les arborescences conduisant à décomposer une catégorie d'individus ou chercher leurs caractéristiques distinctives.

Feuille d'hypothèse :

Pour tracer un modèle sous forme de graphe de relations destiné à être testé.



Propriétés de la vue

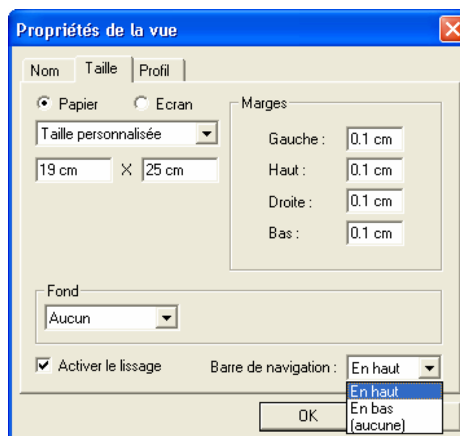
Pour modifier le nom de la vue, son format et la population sur laquelle les résultats seront calculés :

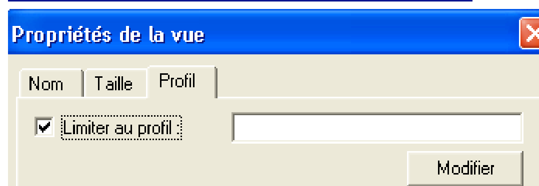
Taille :

Le mode d'affichage est paramétrable avec **Activer le lissage** qui permet de lire par transparence (cartes, AFC…).

Profil :

Pour sélectionner les observations sur lesquelles effectuer le calcul des objets situés dans la vue.

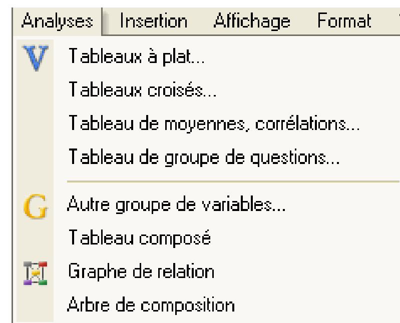




Menu Analyses

Pour définir les tableaux à mettre dans le tableau de bord.

Ces tableaux peuvent également être construits en glissant une variable depuis la fenêtre des variables ou en groupant des variables placées individuellement dans le tableau de bord (Tableau de groupe).



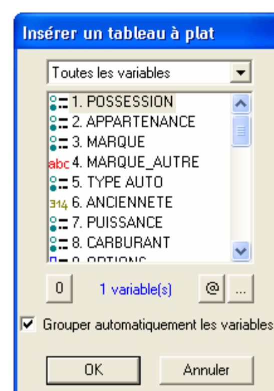
Tableaux à plat

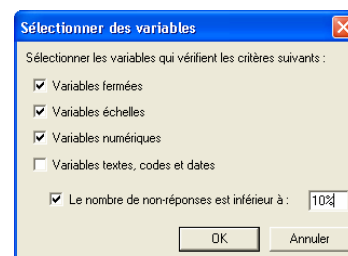
Sélectionner une, plusieurs, ou toutes les variables d'un type donné ou suffisamment documentées (**'...'**).

Utiliser :

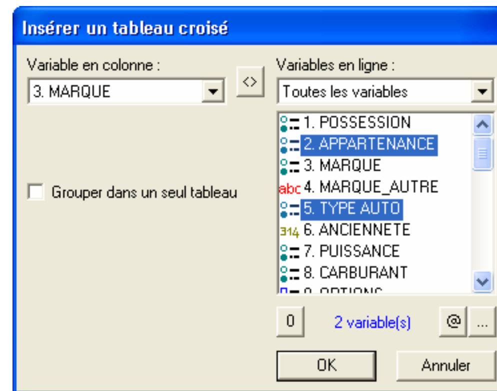
@ pour avoir toutes les variables.

0 pour annuler la sélection.





Tableaux croisés



Pour croiser une variable nominale avec une autre variable. Les numériques sont mises en classes.

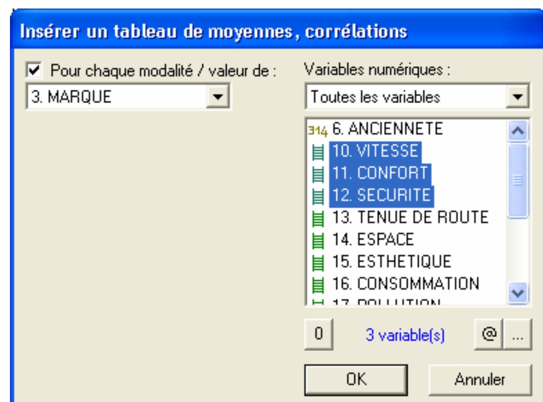
Il faut répéter l'opération d'insertion de tableau croisé pour créer deux tableaux croisés contenant des variables différentes.

Le croisement d'une variable avec plusieurs autres est obtenu en sélectionnant plusieurs variables en ligne.

On obtient autant de tableaux croisés que de variables lignes, ou un seul en cochant **Grouper dans un seul tableau**.

Pour croiser deux variables numériques mises en classes, procéder directement en groupant les deux variables.

Tableaux de moyennes

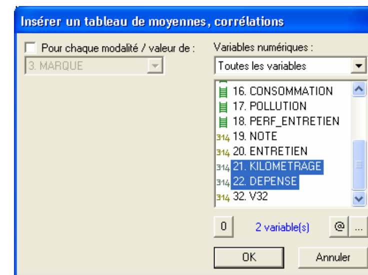


Pour construire un tableau de moyennes croisant une nominale avec des numériques (ou échelles traitées en nombres), cocher **Pour chaque modalité / valeur de** et sélectionner la variable nominale.

Choisir les variables numériques en les cliquant dans la liste proposée.

On obtient un tableau de moyennes croisées. Pour changer le sens du tableau, afficher les **Propriétés des groupes**, et dans l'onglet **Tableau**, indiquer la position des variables numériques en choisissant l'option **En ligne** ou **En colonne** dans l'option **Placement des variables numériques.**

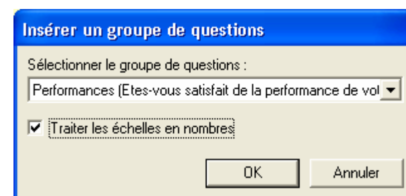
Corrélations



Pour insérer une corrélation et construire un tableau ne comportant que des numériques, sélectionner directement les variables dans la liste de droite.

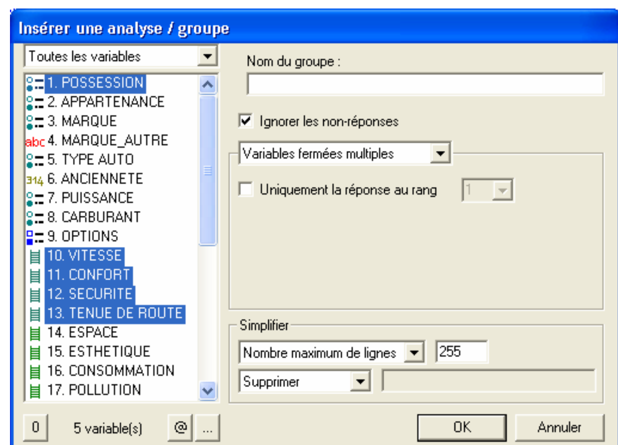
Désélectionnez **Pour chaque modalité / valeur de**.

Tableaux de groupes



Pour construire les tableaux correspondant aux groupes définis dans la liste des questions.

Autre groupe de variables



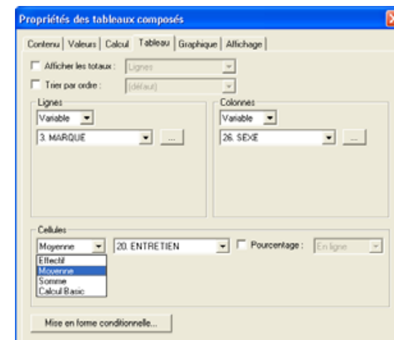
Sélectionner dans la liste de gauche les variables à mettre dans le tableau et préciser les modes de calcul pour chaque type de question présent dans le tableau (échelles traitées en nombres ou non, numériques mises en classes ou non).

Pour la présentation finale du tableau, aller dans les **Propriétés des groupes**, et revenir :

- dans l'onglet **Contenu**, pour fixer le pivot des groupes de nominales en mettant la variable correspondante en tête de liste.

- dans l'onglet **Tableau,** choisir entre **Croiser** et **Accoler** dans l'option **Disposition** pour les groupes de nominales, ou fixer la position des numériques pour les groupes nominales/numériques en choisissant **En ligne** ou **En colonne** dans l'option **Placement des variables numériques.**

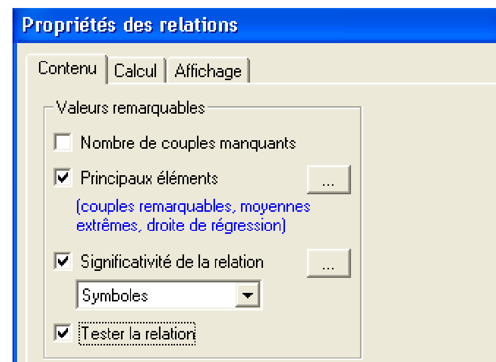
Tableau composé (*)



Introduit un nouvel objet de type tableau composé.

Choisir l'article **Tableau composé** dans le menu **Analyses** et dans les **propriétés du tableau composé**, définir le contenu des lignes, colonnes et cellules du tableau (Voir la partie sur les tableaux composés).

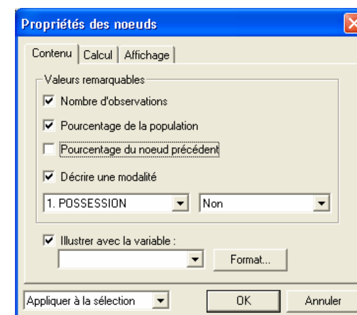
Graphe de relations (*)



Ouvre une nouvelle vue pour représenter les relations entre variables.

Voir Partie Graphe de relations.

Arbre de composition (*)



Ouvre une nouvelle vue avec un nouvel objet de type arbre de composition.

Choisir **Arbre de composition** dans le menu **Analyses**, puis dans le panneau de **propriétés des nœuds**, définir les règles de construction de l'arbre.

Voir Partie Arbres de décision.

Menu Relations (*)

Pour approfondir l'analyse à partir d'un graphe de relations.



Les articles s'appliquent aux relations sélectionnées.

Analyser :

Pour faire apparaître le tableau correspondant.

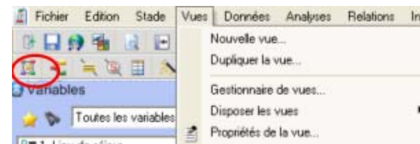
Tracer toutes les relations :

Pour tracer les relations entre les variables sélectionnées.

Grouper :

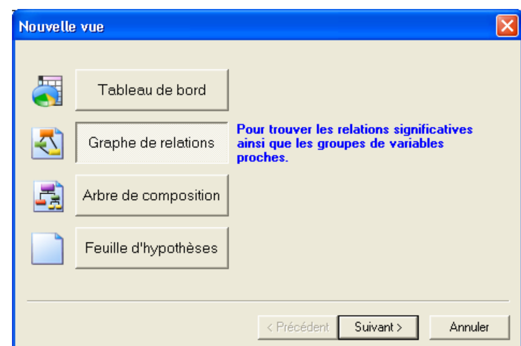
Pour mettre dans un même pavé les variables sélectionnées ou les variables proches (formant un circuit dans le graphe).

Tracer un nouveau graphe (*)



Pour créer un nouveau graphe de relations :

Lancer une nouvelle vue et choisir **Graphe de relations** ou en utilisant le raccourci .



Un graphe peut également être dessiné directement dans la vue en utilisant l'outil relation



situé dans la barre d'outils relative aux relations.

Les relations peuvent également être tracées avec la souris en maintenant la touche **Ctrl** enfoncée.

Choisir un type d’analyse (*)

Après le lancement d’un nouveau graphe de relations, ce dialogue permet de représenter des relations entre variables sous forme de graphe.

Formuler des hypothèses :

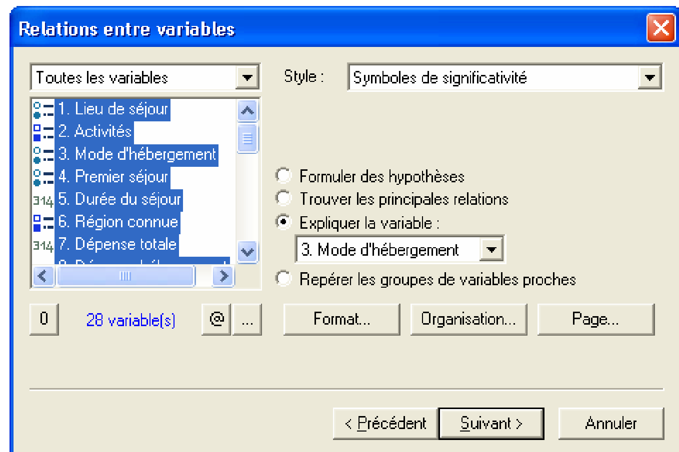
Pour dessiner des relations puis tester le modèle (Menu **Relations** / **Tester**)

Trouver les principales relations :

Pour afficher toutes les relations significatives entre les variables sélectionnées.

Expliquer la variable :

Pour trouver parmi les variables sélectionnées celles qui influencent significativement la variable à expliquer.



Détermination des seuils pour les tests (*)

Les relations sont établies sur la base du test bivarié correspondant aux types des variables de chaque couple.

Jouer sur les seuils et les règles de calcul pour obtenir plus ou moins de relations.

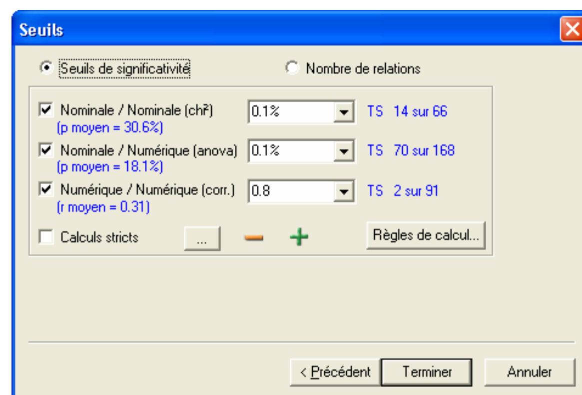
Règles de calcul :

Pour indiquer comment traiter les différents types de variables, et notamment les variables numériques et échelles. Le test appliqué dépend de ce choix.

Calculs stricts :

Pour définir les conditions d’application des tests (effectifs minimum).

Si la recherche se fait sur le nombre de relations, elles apparaissent par significativité décroissante.



Présenter le graphe de relations (*)

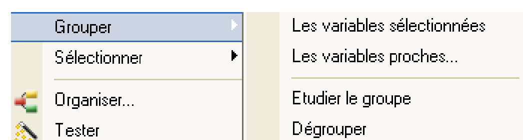


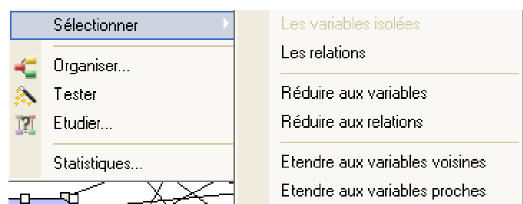
Pour approfondir l'analyse à partir d'un graphe de relations.

Les articles s'appliquent aux relations sélectionnées.

Analyser : pour faire apparaître le tableau correspondant.

Grouper, Sélectionner (*)





Pour mettre dans la même vignette les variables sélectionnées ou les variables proches.

Les variables proches sont reliées par un chemin (circuit dans le graphe).

Dégrouper :

Affiche dans la même vue les variables du groupe.

Etudier le groupe :

Ouvre une nouvelle vue pour l'analyse des variables du groupe.

Sélectionner :

- Sélectionner dans le graphe les relations ou les variables uniquement (et leur appliquer ensuite une action) ;

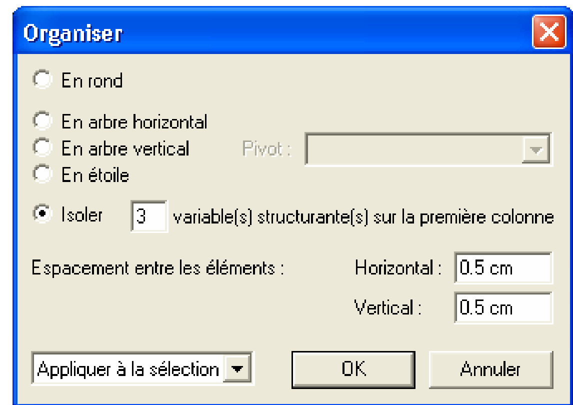
- Rechercher les variables proches d'une variable sélectionnée (directement reliées à elle) ou voisines (reliées par un chemin de plusieurs relations).

Organiser (*)

Pour rendre plus lisibles les graphes denses en les disposant en arbre ou étoile et :

- faire ressortir une variable sélectionnée comme pivot.
- faire ressortir les variables structurantes (celles qui ont le plus de relations avec les autres).

Le raccourci permet d'accéder directement à ce dialogue.

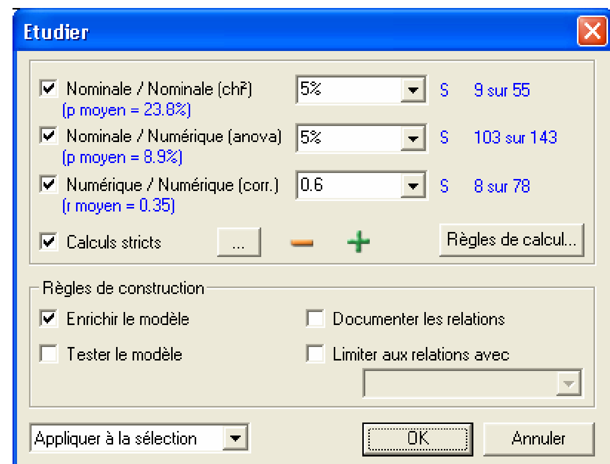


Tester, enrichir (*)

L'article **Tester** du menu **Relations** permet de modifier le nombre ou la nature des relations en jouant sur les seuils et règles de calcul.

Tester le modèle : pour voir la significativité des relations (graphe d'hypothèse notamment).

Enrichir le modèle : pour faire apparaître les relations significatives non envisagées dans le modèle (notamment pour vérifier que les variables explicatives soient indépendantes).



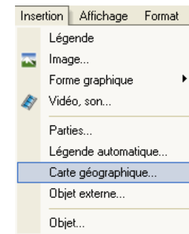
Menu Insertion

Pour documenter les tableaux de bord en introduisant des légendes ou objets graphiques :

- **Légende :** entrer le texte de la légende.

- Image, Forme graphique, Vidéo son… : sélectionner la forme ou le fichier.

- Parties : pour insérer un titre de partie ou l'ensemble des variables de la partie.



Légende automatique

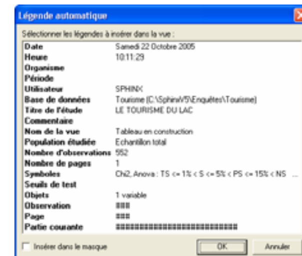
Pour insérer dans une légende des informations définies de manière automatique. Les informations automatiques sont mises entre crochets. Par exemple :

- [Date] / [Heure]

- [Titre de l'étude]

- Nombre de répondants : [Nombre d'observations]

- Page : [Page]



Carte géographique

Pour présenter des données dans une carte. Choisir la carte dans la liste du haut :

- France : Régions, Départements, Départements par régions, Académies. Pour la France des communes, commander la carte du ou des départements désirés.

- Europe, Monde.

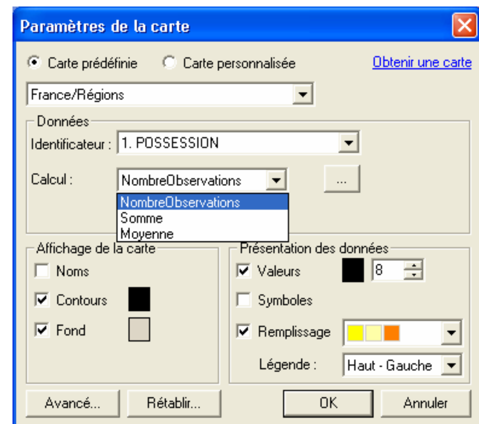
Identificateur : Indiquer la variable contenant la donnée géographique.

Calcul : Définir les éléments statistiques à faire figurer sur la carte :

- **Nombre Observations :** affiche l'effectif de chaque zone de la carte.

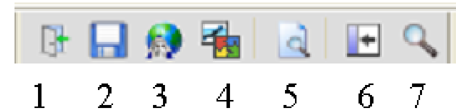
- Somme et **Moyenne** : valeurs d'une autre variable, ou d'un calcul pour les observations de la zone. Choisir la variable ou définir la formule de calcul avec **…**

Affichage de la carte et **Présentation des données :** Pour définir les légendes et codes couleurs.



Icônes et raccourcis de la barre d'outils

Gérer l'environnement et échanger



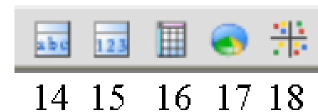
1/ Fermer et revenir au panneau central.
2/ Enregistrer.
3/ Publier sur Internet.
4/ Envoyer dans Office (Word, PowerPoint, Excel).
5/ Aperçu de la vue.
6/ Cacher / Afficher la liste des variables.
7/ Zoom sur la vue.

Variables, groupes, légendes et relations



8/ Sélectionner une zone avec la souris.
9/ Insérer une variable.
10/ Insérer un groupe de variables ou grouper les variables sélectionnées.
11/ Insérer une légende : cliquer sur l'icône puis dans la vue.
12/ Tracer une relation : cliquer sur l'icône puis relier les deux variables avec la souris ou sélectionner les variables puis cliquer sur l'icône.
13/ Supprimer la variable sélectionnée.

Fixer le contenu des variables et des groupes



14/ Afficher / Cacher le texte de la question ou le commentaire.
15/ Afficher / Cacher les valeurs remarquables (taux de réponse, moyenne…).
16/ Afficher / Cacher le tableau.
17/ Afficher / Cacher le graphique.
18/ Afficher / Cacher la carte factorielle.

Conditions d'affichage des variables, groupes, relations et arbres



19/ Placer automatiquement les variables et les groupes (disposition, taille, espacement).

20/ Modifier les propriétés de la vue.

21/ Modifier les propriétés de la variable ou du groupe sélectionné.

22/ Mise en forme automatique des variables, relations et arbres en fonction de leur contenu.

Construire un graphe de relations et fixer son contenu



1/ Construire un nouveau graphe.

2/ Organiser un graphe (disposition en étoile, arbre).

3/ Afficher / Cacher les éléments remarquables de la relation (couples de modalité, valeur maximum ou équation de la droite de régression).

4/ Afficher / Cacher le symbole de significativité du test.

5/ Afficher / Cacher le tableau lié à la relation.

Tester et ajuster le seuil des relations



6/ Tester la relation en relançant le calcul.

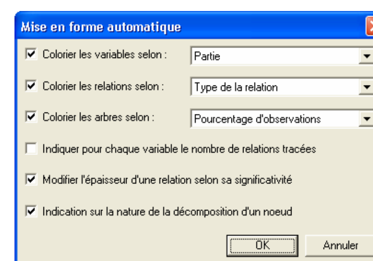
7 – 8 / Réduire / Augmenter le nombre de relations du graphe sélectionné en augmentant/réduisant les seuils.

9/ Afficher le dialogue de paramétrage des tests à appliquer au graphe sélectionné.

10/ Tracer les relations entre les variables.

11/ Insérer un arbre de composition.

Mise en forme automatique



Pour spécifier la nature des variables, relations, arbres.

Colorier les variables selon : Partie, Type de variable, Nombre de non-réponses.

Colorier les relations selon : Significativité de la relation, Type de relation, Signe de corrélation.

Colorier les arbres selon : Pourcentage d'observations, Pourcentage des nœuds précédents, Variable du profil.

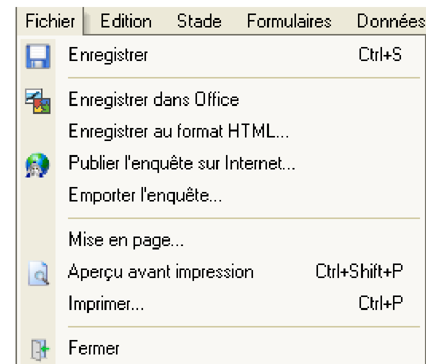
14. Menus communs aux formulaires et aux tableaux de bord

Menu Fichier

Produire le formulaire ou le tableau de bord

Ce menu gère les « sorties » :

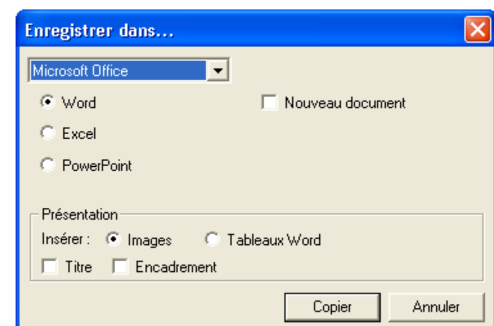
- Enregistrement de sauvegarde : **Enregistrer**.
- Production de documents Word ou PowerPoint : **Enregistrer dans Office**.
- Production de pages html : **Enregistrer au format html**.
- Publication de l'enquête sur Internet : **Publier l'enquête sur Internet**.
- Préparation de l'enquête pour le module opérateur ou un pda : **Emporter l'enquête**.
- **Mise en page** : panneau de propriétés de la vue (formulaire ou tableau de bord).
- **Aperçu avant impression** et **Imprimer** : pour vérifier et lancer l'impression de la vue.



Enregistrer dans Office

Transférer le formulaire courant dans l'application choisie (Word – Excel – PowerPoint).

Transférer le tableau de bord courant dans l'application choisie (Word – Excel – PowerPoint), en indiquant la présentation des éléments transférés (**Images** ou **Tableaux Word**).



Emporter l'enquête (*)

- Répertoire pour le module opérateur :

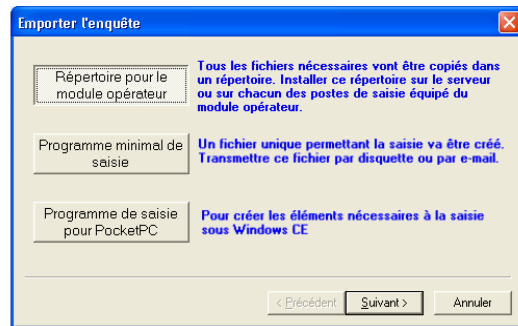
Pour créer les éléments nécessaires à la saisie sur les postes des télé-acteurs et fixer les droits de l'opérateur : saisir, modifier, importer.

- Programme minimal de saisie :

Programme autonome que l'on peut envoyer par mail. Les réponses reviennent dans la boîte électronique indiquée en paramètres. Attention : ce format est souvent interdit par les firewalls anti virus.

- Programme de saisie pour PocketPC :

Créer les éléments nécessaires à la saisie sous Windows CE.

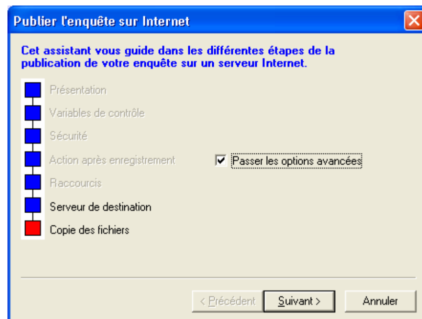


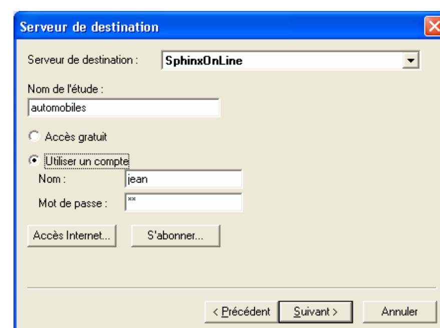
Publier l'enquête sur Internet

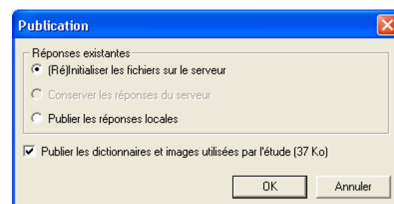
Publication rapide

Un compte de 5 000 unités sur SphinxOnline est ouvert à l'achat de la licence. Il permet de mettre vos questionnaires sur Internet en quelques clics.

Publication gratuite sur le compte public : aucune unité n'est décomptée, mais les données sur ce compte ne sont pas pérennes.







Etude "automobiles"

Vendredi 2 Septembre 2005
SphinxOnLine (public/automobiles)
Ce fichier HTML se trouve dans le répertoire de l'enquête

- http://www.sphinxonline.net/public/automobiles/questionnaire.htm
- http://www.sphinxonline.net/public/automobiles/resultats.htm
- http://www.sphinxonline.net/public/automobiles/tables.htm

Lorsque les fichiers ont été publiés, les liens vers les formulaires, les tableaux de bord ou la table des données sont communiqués.

Assistant de publication (*)

Présentation

L'option Html doit être privilégiée lorsqu'on ne connaît pas le type d'accès des correspondants (débit et type de navigateur).

(*) Variables de Contrôle

Clé : Une variable clé est automatiquement générée. Son contenu est communiqué au répondant après enregistrement. Il pourra ainsi revenir sur ses réponses ou continuer la saisie.

Date de saisie, Adresse IP : La date et l'heure de saisie sont enregistrées ainsi que la référence de la machine qui a répondu.

(*) Sécurité :

Pour empêcher les requêtes non autorisées.

Identificateurs de saisie

Les identifiants de saisie sont communiqués à chaque répondant. Ils leur permettent d'accéder au questionnaire.

(*) Action après l'enregistrement :

Préciser le contenu du bouton destiné à mettre fin à la saisie et à lancer l'enregistrement des réponses.

Définir ce qui advient après cette action (rappel des réponses, proposition d'un lien vers la page d'un site ou par exemple, vers les résultats de l'enquête).

(*) Raccourcis :

Ce dialogue définit les liens donnant accès au questionnaire ou aux résultats.

Ce sont des requêtes au serveur d'enquêtes :

Saisir : créer un nouvel enregistrement.

Consulter : voir le contenu des enregistrements existants.

Modifier : accéder à un enregistrement existant et le modifier, une clé est demandée.

Voir les résultats : afficher un tableau de bord.

Voir les tables : lister les données brutes.

La syntaxe des liens est la suivante :

www.sphinxonline.net/compte/enquête/raccourcis.htm

(*) Serveur de destination

SphinxOnline est le serveur par défaut.

La publication peut également se faire sur un autre serveur équipé du logiciel serveur d'enquêtes.

Après avoir choisi le serveur, on précise le nom utilisé pour publier l'enquête, et le compte sous lequel les fichiers seront enregistrés. Ces informations seront reprises dans les liens vers les questionnaires ou les résultats :

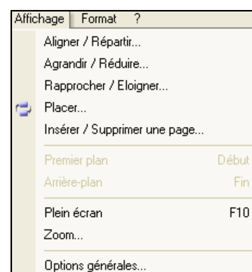
www.serveur/compte/nom_etude/…

(*) Copie des fichiers

En retour, les liens construits à l'étape **Raccourcis** sont envoyés.

Menu Affichage

Ce menu fixe les règles de placement des objets dans la vue (**Aligner/ Répartir, Agrandir / Réduire, Rapprocher / Eloigner**) ainsi que le découpage des vues en page (**Insérer / Supprimer une page**).

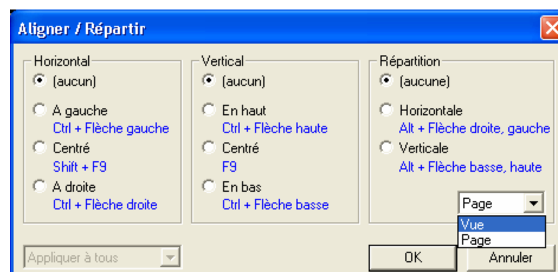


Aligner/ répartir, Agrandir/ Réduire

Aligner /Répartir :

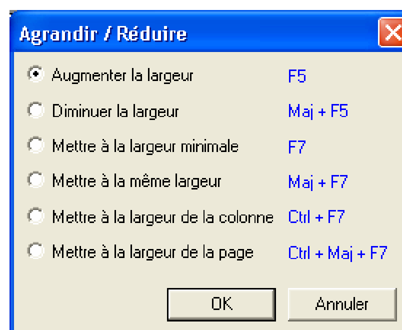
Attention :

La disposition des questions dans les formulaires publiés au format html dépend de leur alignement exact. Les raccourcis **Ctrl+flèche** sont le meilleur moyen pour y parvenir.

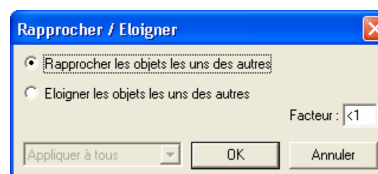


Agrandir / Réduire :

On obtient facilement des formulaires bien présentés en mettant toutes les questions à la même largeur (**Maj+F7**) ou à la largeur de la colonne (**Ctrl+F7**).



Voir les raccourcis dans le menu **? / Raccourcis et Accélérateurs**.

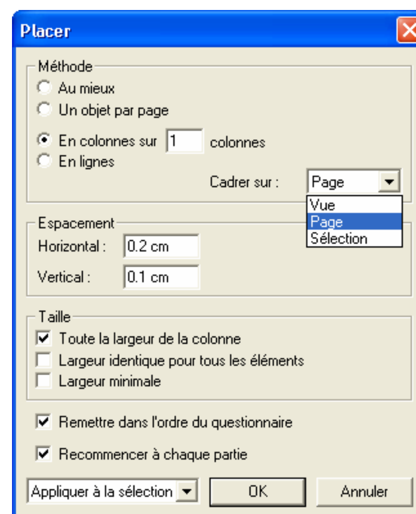


Placer, Insérer

Idéal pour obtenir rapidement une présentation homogène.

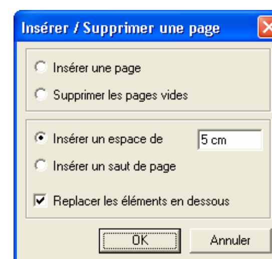
Les espacements sont relatifs à l’espace entre 2 questions (espace vertical et horizontal) et entre colonnes (espace horizontal).

Les paramètres de taille s’appliquent également à tous les types d’objets.



Insérer / Supprimer une page

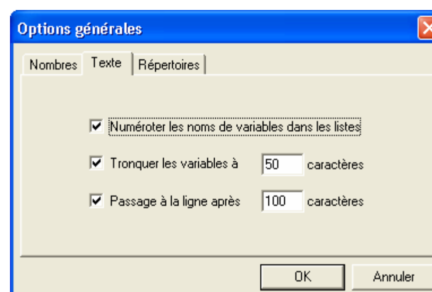
Insérer une page :
Insère une page vide avant la page contenant un objet sélectionné.
Insérer un saut de page :
Fait passer à la page suivante après l’objet sélectionné.



Options générales

Nombres :
Pour fixer le nombre de décimales des nombres et pourcentages.

Texte :
Pour fixer les règles d’affichage des variables, de la numérotation, et de découpage des textes.

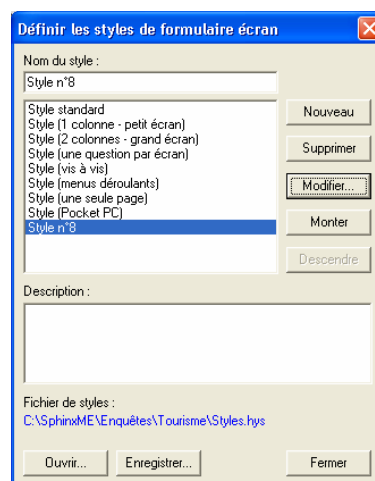
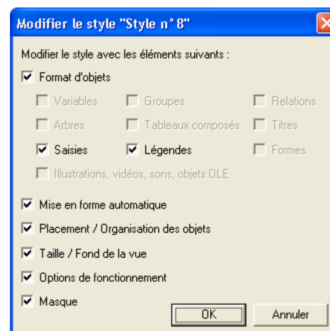


Menu Format

Format

La notion de format renvoie à un ensemble de paramètres définissant les objets d'un formulaire ou d'un tableau de bord. **Copier le format** : copier les paramètres définissant l'objet sélectionné. **Coller le format** : après l'avoir sélectionné, donner à un autre objet, les propriétés de celui dont on vient de copier le format.	

Définir les styles

Un style correspond à un ensemble de formats pour le formulaire et ses objets. A la création d'un nouveau formulaire ou d'un nouveau tableau de bord, les questions ou variables auront les propriétés du style par défaut ou du style courant. **Appliquer un style** : Donner au formulaire ou au tableau de bord les propriétés du style sélectionné. **Définir les styles** : Pour enregistrer un nouveau style correspondant aux propriétés des objets du formulaire ou du tableau de bord actif. Attention : vérifier que le formulaire ou le tableau de bord de référence ne contient pas deux objets du même type, de format différent.	

15. Enquête en ligne et modules complémentaires

Enquête et analyse de données en ligne

Avec le Sphinx, vous pouvez utiliser Internet de 2 manières :

- A la manière du module opérateur en donnant la possibilité à vos interlocuteurs de répondre à vos questionnaires ou de consulter vos tableaux de bord via Internet. Pour cela, à partir de votre logiciel Sphinx Plus² ou Lexica, vous publiez votre enquête sur un compte SphinxOnline.
- A la manière d'une application ASP complète qui vous permet, directement sur Internet, de construire un questionnaire et de faire de l'analyse de données. Pour cela, vous devez lancer l'application ASP de Sphinx. Cela se fait dans l'interface d'un compte SphinxOnline.

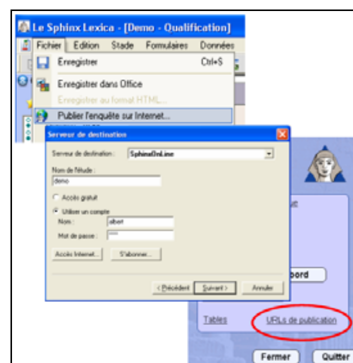
Le compte SphinxOnline à utiliser est celui qui vous a été ouvert à l'achat du logiciel ou auquel vous avez souscrit pour utiliser la version ASP de Sphinx (Sphinx Déclic).

Publier une enquête sur Internet

Vous préparez votre enquête, questionnaire et tableau de bord dans votre logiciel, et vous la publiez depuis l'environnement multimédia :

Fichier / Publier l'enquête sur Internet (raccourci) / Suivant (case **Passer les options avancées** cochée) / Entrez **Compte** et **Mot de passe** / **Terminer.**

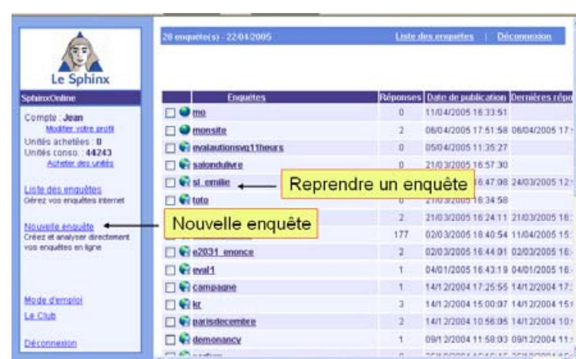
En retour, vous recevez les liens (URL de publication) que vous utiliserez pour diffuser les formulaires et tableaux de bord en les communiquant à vos correspondants.



Ces liens sont accessibles à partir du panneau de commandes central (URL de publication).

Utiliser la version ASP de Sphinx

Le site de Sphinx (www.lesphinx-developpement.fr) donne accès au compte SphinxOnline qui vous permet de créer une nouvelle enquête ou de reprendre une enquête en cours.



L'interface correspond à celle du logiciel et tout se passe sur le web sans avoir à installer un logiciel sur sa machine. Il suffit de disposer d'un navigateur pour :

- définir les questions et mettre en forme le questionnaire,
- diffuser les questionnaires sur Internet ou saisir les réponses au clavier en cas de diffusion papier,
- analyser les résultats.

La table des données ainsi recueillie peut être affichée et recopiée pour être sauvegardée comme fichier texte sur votre poste local.

Vous pouvez également importer ces données directement depuis Sphinx si vous disposez également d'une version locale du logiciel.

Les modules complémentaires

Les applications suivantes font l'objet de modules complémentaires :

- L'utilisation d'un scanner pour automatiser la saisie de questionnaires papiers,
- La mise en place d'un système CATI (Collecte de données assistée par téléphone et informatique).
- L'installation dans un Intranet d'un système privatif d'hébergement d'enquêtes ou de création de questionnaires et d'analyse de donnée en ligne.
- La création dans un Intranet d'une plate-forme d'administration d'enquêtes.

Module scanner

Ce module permet de créer depuis Sphinx des formulaires directement lisibles par un scanner. A cette fin, le logiciel place sur le formulaire les marques qui permettent au scanner de reconnaître automatiquement les informations qu'il contient. Ce module très simple à mettre en œuvre ne se justifie vraiment que pour de gros volumes de saisie.

Système CATI

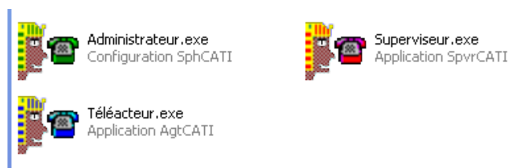
Ce système permet de faire travailler simultanément plusieurs téléacteurs se partageant une base d'appels téléphoniques. Il comporte un module de supervision et d'administration.

Le module de supervision est destiné au gestionnaire du centre d'appels et permet de :

- Choisir un mode d'usage de la téléphonie, composition manuelle, automatique, prédictive…
- Gérer les téléacteurs en les affectant à différentes campagnes menées simultanément.
- Contrôler la productivité des téléacteurs sur la base de ratios de temps et d'efficacité.

Le module d'administration permet de :
- Créer et gérer la base d'appels.
- Créer les scripts d'appel.
- Elaborer les tableaux de bord pour le suivi des résultats de la campagne.

Il est utilisé par la personne responsable de l'enquête.



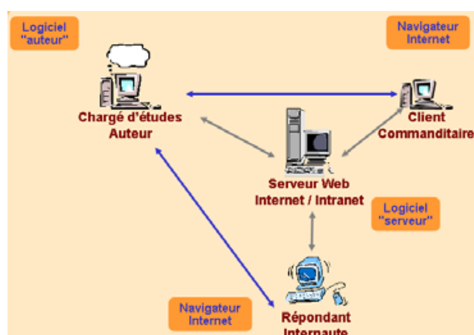
Serveur d'enquêtes

Le serveur d'enquêtes est une version du logiciel qui s'exécute via le web. A partir d'un navigateur web :

- Les répondants font apparaître le questionnaire en cliquant sur un lien Internet.

- Le destinataire de l'étude suit les résultats en utilisant des liens vers les tableaux de bord.

Au préalable le chargé d'études a préparé son enquête avec le Sphinx et publié les fichiers correspondants vers le serveur.



Ce système utilisé pour proposer le service SphinxOnline peut être acquis par les organisations qui souhaitent une appropriation complète de leur application Internet.

L'installation d'un serveur d'enquêtes est complétée par la mise à disposition d'un système de gestion de comptes permettant de gérer les applications de différents services ou utilisateurs.

Peuvent être ajoutées de manière optionnelle :
- la création de questionnaires et d'analyse de données en ligne.
- la gestion d'e-mailing.

Plate forme EvalNet

La plate forme EvalNet permet de gérer au travers d'un Intranet des enquêtes multiples et les publics auxquels elles sont destinées. Le système est entièrement paramétrable et permet de :
- définir les procédures d'inscriptions et les groupes de répondants.
- définir les enquêtes et les groupes de répondants auxquels elles sont destinées.
- gérer les accès et l'avancement des procédures de consultation.

Module multilangue

Le module multi-langue permet de réaliser des enquêtes en différentes langues, et d'agir sur :

- la langue du logiciel utilisée dans les dialogues et menus ;

- la langue du questionnaire, qui sera utilisée pour les chaînes de caractères générées ;

- le paramétrage d'enquêtes multi-langues, dans lesquelles coexistent toutes les traductions.

Méthodologie

16. Rappel des principes de l'enquête par questionnaire

Dans la plupart des cas, l'enquête par sondage est une voie de recherche très efficace pour obtenir des informations. C'est l'outil le plus utilisé dans les études qualitatives et quantitatives, dans les enquêtes et les sondages.

Mais pour être fiable et efficace, cet instrument doit être mené avec précaution. Il est donc nécessaire de déterminer les objectifs de l'étude avant de définir sa mise en œuvre. A cet égard, plusieurs étapes sont alors à envisager.

Définition des objectifs

La première phase d'une étude consiste à définir précisément le problème qui doit être traité. Afin de tirer le meilleur profit des résultats de l'étude, il est indispensable de savoir ce que l'on recherche et d'avoir une idée de l'intérêt de l'étude.

Connaître les données dont on a besoin permet de concevoir un instrument capable de les produire. C'est donc grâce à la détermination préalable des objectifs de l'enquête que des moyens appropriés pourront être définis.

La détermination des échantillons

La réalisation d'un sondage s'effectue sur une partie de la population appelée échantillon. Mais la détermination de l'échantillon doit être calculée afin de fournir des informations et des résultats cohérents avec ceux qui seraient obtenus si toute la population était interrogée. Il faut donc d'abord définir la population mère pour être en mesure d'interroger un échantillon fiable et représentatif.

C'est tout le problème de l'échantillonnage qui consiste à se poser plusieurs questions : Qui interroger ? Combien de personnes ? Quel échantillon retenir et comment le sélectionner (échantillon représentatif ou non, tirage aléatoire ou respect de quotas) ? Dans quelle mesure les résultats de cet échantillon sont-ils fiables ? Etc.

Même si les contraintes budgétaires et les techniques employées fournissent des éléments de réponses, la théorie des sondages nous apporte des enseignements précis fondés sur des hypothèses contraignantes concernant la sélection des échantillons. Néanmoins, le savoir-faire et l'expérience guident souvent fortement la détermination de l'échantillon à interroger et les logiciels Sphinx proposent une aide à l'échantillonnage.

Les différentes formes d'enquêtes

Lorsque les objectifs de l'étude, la population mère et l'échantillon sont fixés, il s'agit de définir les moyens mis en œuvre pour l'étude. Si on envisage que l'enquête sera la méthode d'observation et de recueil des informations, on doit alors opter pour une technique d'enquête.

Or celles-ci sont très diversifiées. Parmi les plus répandues, nous retiendrons l'enquête par voie postale, l'entretien téléphonique, le questionnaire sur Minitel, l'enquête sur Internet et l'entretien en face-à-face. Ces derniers peuvent se dérouler à domicile, sur le lieu de travail, dans la rue, à proximité des points de vente...

Le questionnaire est le seul lien, dans les enquêtes par courrier ou par Internet, entre l'enquêteur et la population interrogée.

Dans le cas des entretiens téléphoniques ou en face-à-face, la communication entre l'interviewé et l'interviewer passe par le dialogue et le questionnaire devient alors un guide d'entretien ou une grille d'observation où il est possible d'enregistrer les réponses.

Le questionnaire

Le questionnaire est donc destiné à capter, dans la population interrogée, les éléments de réponses aux questions que l'on se pose. Il a alors deux objectifs : provoquer une réaction chez les interviewés et servir de support à l'interviewer qui y enregistrera ses informations de façon complète et précise. La rédaction du questionnaire est à soigner dans ce sens.

Le questionnaire incorpore non seulement les questions à poser mais également les plages de réponses. Il est souvent bon d'indiquer un titre ainsi qu'un commentaire évocateur pour la population interrogée.

C'est un instrument très flexible du fait de la grande variété des questions qui peuvent être posées. L'élaboration d'un bon questionnaire requiert une très grande compétence et peut être intégrée dans des logiciels spécialisés en analyse statistique. C'est le cas du Sphinx qui propose une gamme variée de fonctions destinées à la conception du questionnaire.

Le pré-test

La phase de conception d'une enquête s'achève en général par le test d'une enquête pilote qui permet de valider, sur un nombre restreint de personnes, les choix effectués dans le cadre de l'étude.

Ce test permet de voir si le protocole d'étude est réaliste, si le contenu et la forme des questions sont adaptés aux objectifs de l'étude.

C'est aussi souvent l'occasion de découvrir des erreurs grossières et des oublis, ou encore de vérifier la nécessité de chaque question posée et d'écarter éventuellement celles qui ne répondent pas directement aux objectifs de l'étude.

Le test d'une enquête présente donc l'intérêt de rechercher la meilleure adaptation entre les objectifs de l'étude, les moyens alloués et les méthodes choisies.

Le recueil des données

Quoique fastidieuse, cette phase ne présente aucune difficulté particulière, même si selon les techniques d'enquêtes utilisées, les enquêteurs doivent posséder des compétences plus ou moins importantes.

Cette étape nécessite néanmoins une bonne organisation du travail et peut être effectuée dans des logiciels de gestion de données, de traitement d'enquêtes ou d'analyse statistique.

Actuellement, l'amélioration des communications entre ces logiciels permet d'échanger les bases de données très facilement et de reprendre des données existantes comme s'il s'agissait d'informations obtenues par questionnaires. Dans ce domaine, le Sphinx présente une fonction d'importation des données depuis des traitements de textes, des tableurs ou toute base de données externes.

Le dépouillement et l'analyse de données

Une fois les réponses saisies, on s'intéresse aux résultats qui vont ressortir de cette étude. On peut alors distinguer plusieurs niveaux d'analyse : on commencera par le constat des réponses données par les interviewés, c'est-à-dire le dépouillement. Cette phase sera complétée par des calculs ou des tests statistiques et par un approfondissement des analyses pour parvenir aux résultats significatifs de l'enquête, ce qui permettra d'adapter les décisions et les actions aux conclusions de l'étude.

Le dépouillement des résultats donne rapidement un aperçu de l'ensemble des résultats de l'enquête en produisant des tableaux ou graphiques de résultats et des listes de réponses données.

Il est d’abord conseillé de prendre connaissance des résultats des variables considérées indépendamment les unes des autres et de procéder ensuite à la mise en relation de plusieurs variables.

La phase d’analyse permet d’effectuer des tests et des calculs sur les résultats extraits du dépouillement. Elle a pour objectif d’analyser les résultats de façon précise et d’aider à l’interprétation et à la décision.

Il est souvent nécessaire, après analyse, de revenir sur la définition initiale d’une ou plusieurs variables pour modifier et enrichir la base initiale de données. A ce niveau, on peut transformer le contenu d’une variable en procédant à des regroupements ou à la suppression de modalités de réponses, mais on peut également créer ou calculer de nouvelles variables.

La présentation des résultats

La présentation des résultats significatifs de l’enquête est parfois une tâche complexe : elle nécessite de prendre connaissance des analyses de données pour ne sélectionner que les plus caractéristiques et les plus synthétiques.

Les résultats retenus pour le rapport d’étude sont également ceux qui sont susceptibles de conduire aux prises de décisions et aux actions. Il s’agit ensuite de les présenter dans un rapport d’étude qui, par sa mise en page, ses commentaires, ses graphiques..., mettra en valeur les résultats significatifs.

Dans cet objectif, le Sphinx propose des fonctions adaptées qui permettent de présenter un rapport organisé selon un plan de dépouillement, de synthétiser et résumer les résultats dans des tableaux construits à cet effet et complétés par des commentaires générés automatiquement par le Sphinx.

Des fonctions de mise en forme des tableaux et graphiques de résultats permettent également de distinguer les résultats les plus significatifs dans l’ensemble des informations ressortant de l’étude réalisée dans le logiciel. Enfin, les nombreuses possibilités d’échange avec les autres logiciels offrent une grande souplesse lors de la réalisation du rapport d’étude.

La communication

Le travail d'enquête et d'étude s'apparente à un travail de communication :

- communication amont pour la transmission du questionnaire, l'exposé des questions, le recueil des réponses ;

- communication avale pour la diffusion des résultats, le travail d'argumentation et d'aide à la décision.

Les nouvelles technologies offrent de nouvelles opportunités de mise en page, de présentation, d'illustration par des couleurs ou des images, d'interactivité dans les échanges avec les répondants ou le destinataire du rapport. Avec l'évolution des logiciels, le chargé d'études peut ainsi de mieux en mieux maîtriser cet aspect important de son travail.

17. Les différentes formes d'enquêtes

L'observation directe

L'observation directe consiste à mener une observation sans solliciter la participation consciente des personnes observées. Ceci pose bien sûr des questions d'ordre moral : a-t-on le droit de procéder à l'insu de ceux qu'on observe ? Tout dépend de l'usage qui sera fait des informations recueillies.

Il existe aussi des obstacles d'ordre pratique. En effet, beaucoup d'informations sont inaccessibles par cette méthode. D'autre part, les dispositifs concrets permettant d'assurer ce type d'observation (camouflage, glace sans tain, caméra vidéo) sont coûteux et difficiles à mettre en œuvre. Notons cependant les nombreuses possibilités offertes par Internet. L'analyse des traces (origine, pages visitées, temps passé, clic...) laissées par l'internaute est une modalité de l'observation directe.

Entretien en face-à-face

Les protagonistes de l'entretien se font face et peuvent ainsi dialoguer en utilisant toutes les ressources de la communication interpersonnelle. Les circonstances de ce type d'entretien - communication de sujet à sujet - présentent des avantages certains. L'enquêteur sollicite activement le répondant tout en interagissant avec lui pour réguler l'entretien dans sa durée. Des questions peuvent être précisées ou expliquées, l'interprétation des réponses peut être vérifiée, au risque cependant d'influencer ou de biaiser l'observation. Cette méthode n'est pas exclusive de l'observation directe. L'enquêteur peut, en cours ou en fin d'entretien, noter les caractéristiques du comportement de son interlocuteur. Durée de l'entretien, perception de l'assurance, de la sincérité de l'interlocuteur, présence ou absence de certains indices sur les lieux de l'interview, de comportement a priori définis.

Entretien téléphonique

C'est une autre forme d'entretien. La communication y dispose de moins de ressources. Les protagonistes ne se voient pas, l'enquêteur ignore le cadre dans lequel se trouve le répondant. L'interaction reste possible, mais la bonne compréhension de l'interlocuteur est privée des informations gestuelles. La pression du temps s'exerce différemment.

Enquête par voie postale

Le questionnaire est, dans ce cas, l'unique lien entre l'observateur et la population. Le répondant est seul, libre de répondre ou non, dans l'ordre qui lui convient, sans subir d'autre influence que celle des indications et questions du questionnaire. Il a tout le temps de réfléchir à ses réponses. L'observateur s'est exprimé une fois pour toutes en élaborant des questions qu'il ne peut plus ni modifier ni expliquer. De même n'a-t-il aucun recours auprès du répondant pour vérifier le sens de ses réponses.

Enquête via Internet

Ce type d'enquête se développe avec l'usage de l'Internet. Le questionnaire est accessible sur un site, le répondant lit les réponses sur son écran et entre ses réponses. L'avantage de ce procédé est de supprimer en aval la saisie informatique. Ce moyen permet aussi de gérer la séquence des questions. Une nouvelle question n'apparaît à l'écran que lorsque la question précédente a reçu une réponse. C'est un avantage par rapport aux enquêtes postales dans lesquelles on ne peut dévoiler progressivement les questions. Cette approche est de plus en plus fréquente pour la consultation de panels. Elle reste encore limitée pour le grand public par le faible taux de connexion des ménages.

Enquête en laboratoire et panel

La situation expérimentale consiste à mettre l'individu dans un contexte contrôlé par l'expérimentateur. Il est possible ainsi, en construisant des plans d'expérience, d'isoler les effets de chacune des actions envisagées.

Avec les panels, on professionnalise l'échantillon en recrutant dans une population considérée, des individus qui acceptent de répondre aux consultations dont ils seront l'objet. Ils sont en général rémunérés et formés au rôle qui est le leur : répondre le plus objectivement possible aux questions qui leur sont périodiquement posées par l'institut qui gère le panel.

	AVANTAGES	INCONVENIENTS
Enquête par observation directe	- Objectivité dans l'observation des faits ou comportements. - Perturbation minimum du fait de l'enquêteur.	- Impossibilité d'observer des opinions ou attitudes. - Difficulté de mise en œuvre pratique (condition de l'observation, formation de l'enquêteur). - Problème déontologique. On observe des gens à leur insu.
Enquête en face à face	- Permet l'observation des attitudes et comportements. - Bon contrôle de l'échantillon sondé : les personnes contactées sont « contraintes » de répondre. - Possibilité de dévoiler progressivement les objectifs de l'enquête. - Souplesse liée à l'enquêteur : adaptation du vocabulaire, interprétation des réponses, précisions apportées. - Possibilité d'entretiens plus longs.	- Coûteux. - L'enquêteur influence le répondant. - Tout dépend de la qualité des enquêteurs, de leur formation à l'enquête, et de leur sérieux sur le terrain.
Enquête téléphonique	- Moins coûteux que face-à-face. - Moins d'influence liée à l'enquêteur.	- Difficulté à poser correctement des questions à réponses assistées. - Impossibilité de passer des questionnaires trop longs.
Enquête par courrier postal ou électronique	- Coût moindre surtout avec Internet. - Le répondant ne subit pas l'influence de l'enquêteur. - Le répondant a le temps de la réflexion, ce qui permet une meilleure approche des questions d'opinion.	- Faible taux de réponse. - Absence de contrôle a priori de l'échantillon. - Forte influence liée au questionnaire et à sa logique.

18. Un modèle pour concevoir un bon questionnaire

Celui qui rédige un questionnaire peut toujours ramener les questions qu'il envisage à l'un des 4 grands thèmes suivants.

Ceux-ci peuvent être étudiés indépendamment les uns des autres, mais la richesse de l'enquête naîtra de la manière dont on est capable de les relier dans un système.

Les grands thèmes d'une enquête

Les quatre grands thèmes suivants peuvent s'appliquer à l'étude de tout type de population. Le 4ème thème ne concerne que les populations humaines.

- **Identité** : qui interroge-t-on? Quels objets observe-t-on?

- **Comportement** : Que font ceux qu'on interroge, comment agissent-ils? Quelles sont les propriétés des objets observés ?

- **Motifs contraintes**: quelles sont les raisons qui guident les comportements, expliquent les actions ? A quelles contraintes, mécanismes sont soumis les objets étudiés ?

- **Opinions et valeurs** : quelle signification les sujets accordent-ils à leur comportements, sur quelles valeurs se fondent leurs motifs d'action ?

Concevoir le questionnaire comme un système

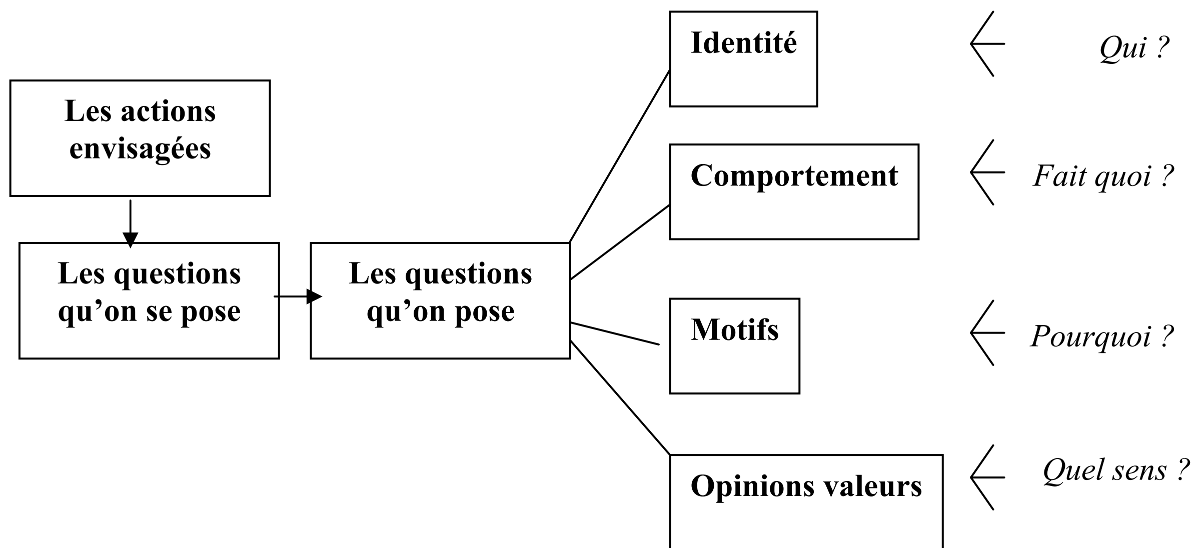
Les thèmes qui structurent le questionnaire peuvent être envisagés comme un système situant les questions les unes par rapport aux autres. Ainsi, l'explication d'un comportement peut être recherchée dans des facteurs d'identité suivant les modèles du déterminisme social ou dans la prise en considération des motifs en référence au modèle de décision rationnelle. Toutes les relations envisageables entre les différents thèmes peuvent faire sens en renvoyant aux grandes théories du domaine étudié.

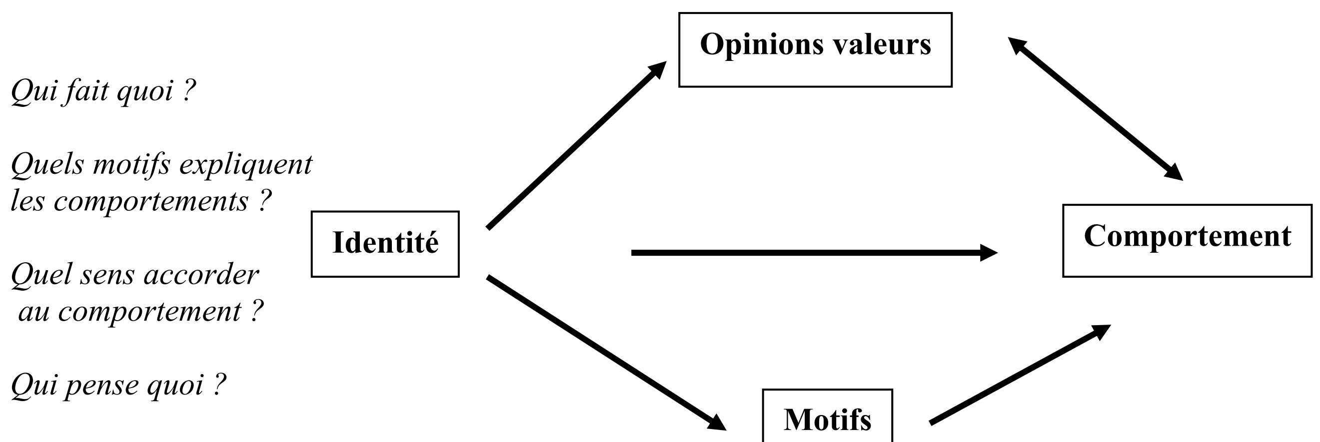
Raisonner ainsi sur les relations entre les questions permet d'affiner le questionnaire en anticipant sur les analyses qui se révéleront utiles au moment de l'analyse des résultats.

Utiliser le modèle identité, comportement, motifs, opinion et valeurs

Quel que soit votre domaine d'étude, il vous sera utile de revenir sur votre questionnaire en l'analysant du point de vue de ce modèle. Tous les thèmes sont-ils abordés, les manques correspondent-ils à un choix délibéré ou à un oubli ?

Quelles sont les relations envisageables entre questions ? A quel type de théorie renvoient-elles ? ....


Le questionnaire comme liste
Les actions envisagées
Les questions qu'on se pose
Les questions qu'on pose
Identité
Comportement
Motifs
Opinions valeurs
Qui ?
Fait quoi ?
Pourquoi ?
Quel sens ?


Le questionnaire comme système
Qui fait quoi ?
Quels motifs expliquent les comportements ?
Quel sens accorder au comportement ?
Qui pense quoi ?
Identité
Opinions valeurs
Comportement
Motifs

19. Des questions qu'on se pose aux questions qu'on pose

Pièce maîtresse de toute enquête, le questionnaire est à l'articulation du modèle conceptuel, expression des hypothèses et du mode opératoire, instrument d'observation et de recueil de l'information.

De la théorie à l'observation, du concept à la variable, le questionnaire organise, par l'intermédiaire des questions posées et des réponses obtenues, la confrontation des idées aux phénomènes du terrain.

Quelle que soit la sophistication des traitements effectués et le sérieux des chiffres, c'est sur le sens des mots que reposent finalement les contributions de l'étude.

La question qu'on se pose

Toute question posée dans un questionnaire renvoie à une ou plusieurs questions que s'est posée celui qui fait l'étude.

Ainsi, derrière toute question qu'on pose, il y a une question qu'on se pose : une hypothèse. Celle-ci renvoie à une ou plusieurs théories, connaissances préalables dégagées dans une phase préliminaire d'étude documentaire et de réflexion. La qualité du questionnaire repose sur la clarté de ce travail initial.

La question qu'on pose

Sa fonction est de susciter une réponse, donc de provoquer une réaction. Mais pas n'importe quelle réaction. On recherche en fait la réponse à la question qu'on se pose et on veut la connaître avec une objectivité maximale. Il faut donc perturber le moins possible l'authenticité de ce qu'est celui qu'on interroge, ce qu'il fait, sait, ressent, pense et à tout prix, éviter de provoquer une réponse qui serait influencée par les circonstances de l'interrogation.

Neutralité, objectivité, mais aussi clarté. Que le sens de la question soit le même pour celui qui la pose que pour celui qui l'entend. La qualité des réponses dépend de celle de la compréhension entre questionneur et répondant.

La réponse qu'on enregistre

La finesse de l'observation dépend du procédé d'enregistrement de la réponse. Répondre en choisissant parmi une liste de modalités prédéfinie fait perdre la variété et les nuances que permet l'enregistrement d'une réponse librement formulée. Il en va différemment lorsqu'il s'agit d'une grandeur ou d'un nombre. Il suffit alors d'enregistrer tel quel le chiffre annoncé pour saisir dans l'unité considérée toutes les nuances de la réponse. Dans tous les autres cas, la mesure dépend de l'étalonnage de l'instrument.

Choix de l'unité, définition a priori d'un système de codification : dès la conception du questionnaire, il faut imaginer les réponses. Ce travail suppose une connaissance a priori sur les phénomènes abordés. Sans hypothèses, pas d'observation.

Rédiger un bon questionnaire

Réaliser un bon questionnaire nécessite de la méthode et beaucoup de métier.

La méthode : respecter les étapes de la démarche :

Formuler des hypothèses claires, les questions qu'on se pose.

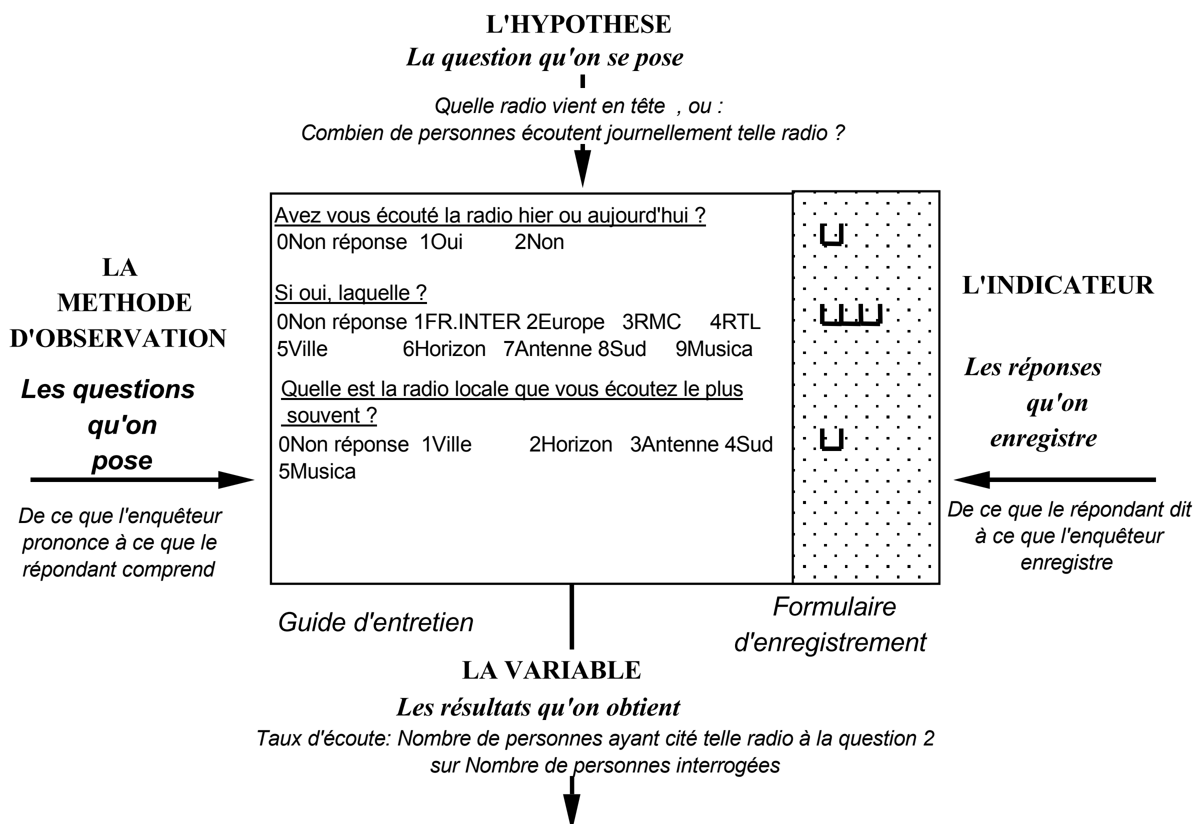
Traduire ces hypothèses sous forme de questions à poser en tenant compte des caractéristiques des populations interrogées, de leur langage et des circonstances de l'interrogation.

Tester le questionnaire en l'administrant en situation. Cette phase fait ressortir les problèmes de langage, les difficultés de compréhension, elle permet d'évaluer la durée de l'entretien...

Dépouiller et analyser les réponses obtenues au cours du test. On pourra ainsi constater que telle question n'est pas vraiment utile, que telle autre manque, que la manière d'enregistrer les réponses n'est pas la plus pratique. On est ainsi conduit à tenir compte, dès la conception du questionnaire, des traitements que l'on souhaite faire.

Le métier : il s'acquiert avec l'expérience. Formuler un bon questionnaire est une tâche difficile qui exige de trouver le meilleur compromis entre les impératifs de la conceptualisation, du terrain et des méthodes statistiques. A cet égard, si vous êtes néophyte, vous gagnerez beaucoup de temps et obtiendrez de bien meilleurs résultats en faisant appel aux conseils d'experts. C'est ce que nous vous proposons au travers du service ***Le Sphinx en direct***. Consultez-nous pour bénéficier de notre expérience.

Le questionnaire est au centre du dispositif d'enquête



20. La manière de poser les questions

Entretien libre : Question ouverte, réponse spontanée

Une question est à réponse spontanée lorsqu'aucune indication n'est apportée au répondant sur la manière de formuler sa réponse. Même s'il s'agit d'une question à réponse fermée - dont la réponse sera codée par rapport à des modalités préétablies - la nomenclature préconçue ne lui est pas communiquée. Par exemple, la profession est demandée sans que la classification en CSP ne soit fournie. Le répondant est donc entièrement libre de sa réponse.

Entretien semi ouvert : Question ouverte, réponse fermée

La réponse librement obtenue n'est pas nécessairement enregistrée telle quelle. L'enquêteur peut interpréter la réponse pendant l'entretien, ou à réception des questionnaires expédiés par courrier. Il applique pour ce faire la codification dérivée des modalités de réponses définies avec la question.

Procéder ainsi suppose :

En face-à-face, que l'enquêteur prenne garde de ne pas dévoiler la liste des modalités qu'il a sous les yeux ;

Par courrier ou Internet, que les modalités de réponses aux questions qui doivent rester ouvertes, ne figurent pas sur le questionnaire expédié, bien qu'elles aient été préconçues et doivent servir ultérieurement à la codification.

Comme le montre le tableau ci-contre, l'entretien est dit ouvert si la réponse est libre et enregistrée telle qu'elle est donnée. Il est semi-ouvert si la réponse est libre, mais la réponse enregistrée suivant la codification préétablie.

Entretien fermé : Question assistée

La question est fermée ou à réponse assistée si la liste des modalités de réponses est communiquée au répondant. On l'aide à répondre en lui proposant de choisir parmi une liste. Dans les enquêtes par courrier ou Internet, les questions sont généralement assistées. Très couramment utilisé, ce procédé simplifie considérablement le traitement. On évite ainsi tout le travail d'interprétation engendré par le système semi-ouvert.

Dans les enquêtes en face-à-face, l'enquêteur peut énumérer la liste, la faire lire sur son document ou utiliser des panneaux écrits en grosses lettres à cet effet. Cette dernière méthode est de loin la meilleure. Au téléphone, l'énumération est le seul moyen d'assister la question. Cela ne va pas sans poser quelques difficultés si la liste est trop longue.

Procéder ainsi a l'avantage de la facilité pour ceux qui font l'enquête. Mais cette méthode comporte également des risques. Le répondant est fortement poussé à répondre. Il faut donc bien préciser que la réponse n'est pas obligatoire et prévoir une rubrique "autre précisez"

L'ordre dans lequel sont énumérées les questions et les modalités de réponses a une influence sur le choix. Lorsque la liste est longue, les dernières modalités sont survolées ou trop rapidement énoncées par l'enquêteur. L'attention se fixe alors sur les premières citations. Si de plus, celles-ci sont des réponses évidentes, le phénomène est amplifié : il faut donc éviter de placer ces modalités en tête de liste. Pour éviter les autres effets d'ancrage sur le début de la liste, il peut être intéressant de soumettre l'échantillon à des énumérations différentes obtenues par permutation. On neutralise ainsi les éventuelles influences en faisant varier la position des termes.

La manière de poser une question peut en affecter le sens

Suivant que la question est assistée ou non, l'information recueillie n'a pas la même signification.

Notons pour l'instant que cette décision affecte le fond de l'enquête et ne doit pas être prise à la légère. Son application correcte par les enquêteurs doit donc être soigneusement contrôlée.

Poser les questions et enregistrer

Manière d'enregistrer la réponse.	**Manière de poser la question.** Question Ouverte: à réponse spontanée.	Question Fermée à réponse assistée.
Fermée Réponses selon Modalités prédéterminées.	**I ENTRETIEN SEMI-OUVERT** Question: Possédez vous une automobile, si oui quelle en est la marque ? Réponse du répondant Citroën Enregistrement réponse 2 0 Non réponse 1Non 2Française 3Allemande 4Italienne 5Japon 6Améric. 7Autre	**II ENTRETIEN FERME** Question: Parmi les qualités`suivantes : confort sécurité, puissance, robustesse, vitesse, faible consommation, faible coût achat quelles sont les 3 plus importantes à vos yeux? Réponse du répondant: Confort, Vitesse Enregistrement réponse 1 5 0 Non réponse 1Confort 2Sécurité 3Puissance 4Robustes. 5Vitesse 6Consomm. 7Prix 8Autre
Ouverte: Réponse telle qu'elle.	**III ENTRETIEN OUVERT** Question: Combien de kilomètres parcourez vous dans l'année avec votre automobile Réponse du répondant 25400 Enregistrement réponse 25400 ou: Question: Possédez vous une automobile, si oui quelle en est la marque ? Réponse du répondant Citroën Enregistrement réponse *Citroën*	**IV** *Le répondant ne peut se satisfaire des réponses proposées. Préciser la catégorie "Autre".*

21. L'art du questionnaire

L'art du questionnaire : la logique de l'entretien

Quelques principes fondamentaux doivent être scrupuleusement respectés.

Introduire le questionnaire

Courrier ou media électronique

Un texte bref situe l'organisme qui réalise l'enquête et donne quelques indications sur les buts poursuivis ainsi que sur la manière de remplir le questionnaire.

Il est important de prévoir une incitation à répondre : explication des buts de l'enquête, proposition de transmettre les résultats, ou cadeau. Tout dépend du budget mais la force de l'incitation peut affecter beaucoup le taux de réponse.

Plus le questionnaire est long, plus faible est l'incitation à répondre. Il doit être clairement rédigé, aéré et occuper un nombre de pages le plus réduit possible.

Les possibilités graphiques et d'animation des médias électroniques peuvent être utilisés comme des incitations à répondre. Mais attention à ne pas surcharger les pages et allonger ainsi le temps de réponse.

Face-à-face

Tout le processus repose sur la qualité des relations que l'enquêteur parvient à établir.

Sa mise, son expression, ses attitudes doivent être adaptées au public qu'il interroge.

Ces paramètres sont à adapter aux circonstances de l'enquête (dans la rue, à domicile...).

Téléphone

Il n'y a pas de différence fondamentale entre la prise de contact en face-à-face et au téléphone mais le téléphone exige une concision et une clarté d'expression encore plus grande.

Si le répondant se déclare indisponible, il est plus aisé d'obtenir un rendez-vous téléphonique. Le simple fait de le solliciter permet parfois de faire tomber l'objection.

Respecter les usages de la conversation

Un entretien a sa logique propre. Comme une conversation, il évolue de propos généraux vers des questions plus précises. En face-à-face ou au téléphone, il faut en tenir compte.

Les questions gênantes ou difficiles doivent être reportées en fin d'interview. Il faut éviter de faire débuter un questionnaire en face-à-face par des questions d'identité. Ce qui se justifie du point de vue de l'analyse risquerait de transformer l'entretien en interrogatoire de police. Il faut prévoir des questions d'introduction ou de transition. Peut être inutiles pour l'analyse, elles ont pour but de rendre l'entretien plus facile.

Enfin, l'ordre des questions doit tenir compte des effets perturbateurs de l'entretien. Les questions à réponse spontanée doivent précéder les questions assistées. Dans le même esprit, il faut révéler le plus tard possible le but précis du questionnaire. En effet, certaines réponses risquent d'être influencées par celui-ci. Même par courrier, il faut respecter cette règle. On veillera également à ce que les questions dont les réponses peuvent être liées figurent sur des pages différentes.

Longueur du questionnaire

Plus un questionnaire est long, plus il est difficile à administrer. Cette contrainte joue en particulier pour les enquêtes par courrier et téléphoniques.

Eviter les questionnaires organigrammes

Il faut éviter de multiplier les questions-renvois. Elles compliquent beaucoup l'entretien et ne se justifient pas toujours. On risque en effet d'imposer au répondant une logique qui n'est pas la sienne et de biaiser ainsi l'observation. Il faut donc se garder de construire un questionnaire comme on conçoit un organigramme et limiter les questions-renvois aux impossibilités strictes de répondre.

L'art du questionnaire : la logique de l'analyse

La difficulté consiste à rédiger un questionnaire qui passe bien et qui permette par la suite les analyses les plus riches possibles.

Ne poser que des questions utiles

C'est une évidence à rappeler car la rédaction fait souvent oublier les objectifs initiaux. En confrontant la liste des questions que se pose le demandeur à celle des questions rédigées dans le questionnaire, il faut vérifier que rien n'a été oublié et que tout est nécessaire. Deux cas peuvent alors se présenter :

Une question du demandeur de l'étude reste sans réponse. Il faut compléter le questionnaire ou constater qu'il est impossible de répondre sérieusement à l'objectif fixé.

Une question du questionnaire ne peut être rattachée à aucune question du demandeur. Il faut la supprimer ou ajouter aux objectifs telle contribution initialement non prévue. Mais assurons-nous alors que c'est une connaissance susceptible d'affecter l'action du demandeur.

Adapter le questionnaire aux analyses projetées

Nomenclatures et analyse par strate

La qualité des résultats est fonction de la taille de l'échantillon. Ainsi, découper un petit échantillon en strates trop nombreuses conduit à des résultats sans signification.

Il faut par conséquent adapter à la taille de l'échantillon, les modalités des questions définissant les strates. Si N est l'effectif total et si le nombre de modalités dépasse N/30 ces modalités définiront au moins une strate non exploitable.

Ainsi, utiliser une nomenclature de C.S.P. en 12 postes en n'interrogeant que 200 personnes nous obligera à regrouper des catégories entre elles pour obtenir des strates significatives.

Type de variables

Suivant le type de questions, les réponses obtenues qualifient des états possibles (variables qualitatives), ou mesurent des grandeurs (variables quantitatives).

Les variables quantitatives sont plus riches et autorisent des traitements plus complexes : calculs de moyennes, d'écarts types, corrélations, régressions, analyse en composantes principales.

Ces indicateurs ou méthodes permettent des analyses plus fines et plus puissantes que celles autorisées par l'usage des variables qualitatives.

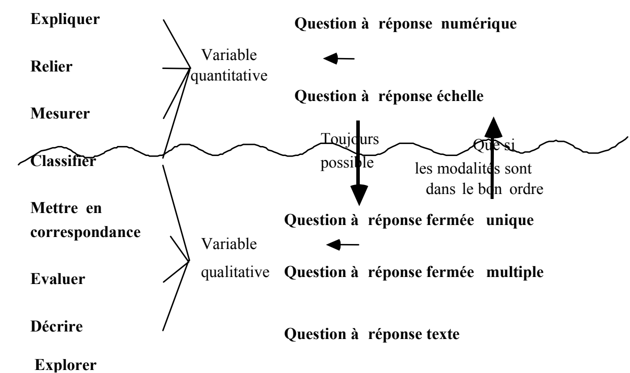
Il faut donc être capable, dès la conception du questionnaire, de définir les traitements envisagés, pour recueillir les informations nécessaires selon le bon format.

Ce choix consiste à déterminer le format de la réponse (texte, codée, fermée, échelle, numérique) en fonction des analyses envisagées.

Dès la conception du questionnaire, tenir compte de l'analyse des données

Certains choix effectués au moment de la rédaction du questionnaire conditionnent fortement les possibilités ultérieures d'analyse.

Ainsi, s'il est toujours possible de transformer une réponse numérique en variable qualitative, l'inverse n'est pas toujours possible. Si d'autre part, aucune question ouverte n'a été prévue, le questionnaire ne permettra aucune véritable découverte, et sans numérique, aucune mesure n'est facilement réalisable.



22. Théorie des sondages et estimation statistique

Réaliser un sondage, c'est substituer à l'étude d'une population entière (la population mère), l'observation d'une partie de cette population, l'échantillon. La théorie des sondages permet :

- de définir les conditions selon lesquelles on peut estimer les propriétés de la population totale à partir des observations de l'échantillon.

- de qualifier l'estimation en indiquant le degré d'erreur ou de risque qu'elle comporte.

Le sondage aléatoire

L'estimation statistique ne peut être effectuée que sous les conditions d'un sondage aléatoire. Celles-ci impliquent que la population soit de très grande taille par rapport à celle de l'échantillon et que chaque individu de la population ait exactement la même chance de faire partie de l'échantillon. Sous ces conditions, le calcul de probabilité montre qu'on obtient un échantillon dont la composition est voisine de celle de la population mère.

L'estimation statistique

Dans l'hypothèse du sondage aléatoire, on peut calculer, à partir d'un résultat observé dans l'échantillon, l'intervalle dans lequel doit normalement se situer la valeur correspondante dans la population totale. Cette fourchette appelée intervalle de confiance indique ainsi la marge d'imprécision que comporte toute estimation. On a l'habitude de la distinguer du risque d'erreur pris en acceptant cette fourchette comme valide. Plus on souhaite réduire le risque, plus la fourchette sera large et les résultats imprécis, au contraire, on peut désirer afficher des résultats plus précis mais avec un risque d'erreur plus grand.

La seule manière d'améliorer la précision sans augmenter le risque d'erreur est d'augmenter la taille de l'échantillon.

Attention: la qualité de l'estimation ne dépend que de la taille de l'échantillon, sous réserve que le tirage est bien aléatoire. Si l'interrogation porte sur une population entière, les résultats sont exacts et il n'y a plus lieu de parler d'estimation.

Si l'exhaustivité n'est pas atteinte, quelle que soit l'importance du taux de réponse, l'estimation n'est possible que si les réponses obtenues sont le fait du hasard. Sa qualité ne dépend que du nombre de répondants.

Les paramètres influençant la qualité d'une estimation

L'intervalle de confiance dépend essentiellement de la taille n de l'échantillon. Par exemple, pour l'estimation d'une proportion p, on le calcule en application de la formule ci-contre. Il est important de constater que l'intervalle de confiance décroît avec la racine carrée de la taille de l'échantillon, ce qui signifie que plus l'échantillon est grand, plus le gain en précision sera faible.

D'autre part, le produit p*(1-p) est de valeur maximum quand p est égal à 0,5 ; ce qui signifie qu'il sera beaucoup plus difficile d'estimer la victoire d'un candidat de deuxième tour au soir des élections (il faudra examiner près de 3000 bulletins) que l'élimination d'un petit candidat au premier tour. Une centaine de bulletins suffisent pour prévoir l'échec d'un candidat rassemblant 10% des suffrages.

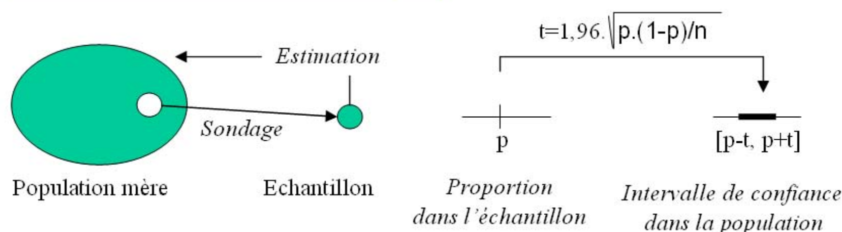
La pratique des sondages

Les conditions théoriques du sondage aléatoire sont très difficiles à réaliser pratiquement. Aucun protocole concret ne permet d'assurer la stricte équivalence des chances d'être interrogé. Même si les numéros de téléphone sont tirés au hasard, la présence ou l'absence au moment de l'appel introduit un biais lié au mode de vie...

Pour cette raison, il est toujours utile de vérifier la qualité d'un échantillon en contrôlant, sur des caractères connus dans la population totale, que les résultats sur l'échantillon sont conformes. Sinon, on dit que l'échantillon est biaisé. Plutôt que d'avoir à le redresser a posteriori, on peut fixer un plan de sondage par quota pour assurer a priori la proportionnalité de l'échantillon. Les limites de cette méthode tiennent à la connaissance de la population à interroger et aux possibilités pratiques de recueillir des réponses à partir de plans de sondage multi-critères très fins. Interroger tant d'hommes, ouvriers, de plus de 50 ans...

Dans la pratique, on combine souvent la méthode des quotas avec une procédure libre pour trouver les individus correspondant aux quotas prédéfinis.

Tirage aléatoire et estimation statistique



Si le sondage est aléatoire, alors on peut estimer à partir de l'échantillon les propriétés de la population totale

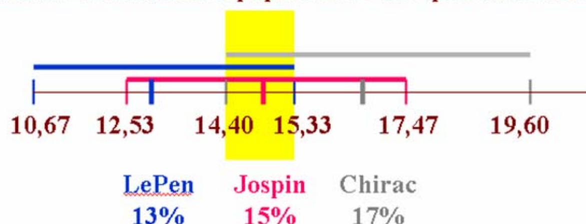
Le 21 avril 2002 n'est pas une surprise statistique

Résultats dans l'échantillon N=800

- Chirac 17 % [14,4 - 19,60]
- Jospin 15 % [12,53 - 17,47]
- Le Pen 13% [10,67 - 15,33]

- N=2000
- [15,35 18,65]
- [13,44 16,56]
- [11,53 14,47]

Estimation concernant la population : tout peut intervenir !



Une application de la notion d'intervalle de confiance :

Aucun des sondages de l'époque ne permettaient en toute rigueur de pronostiquer qui serait présent au deuxième tour.

23. Définir le bon échantillon

La théorie des sondages nous apprend que l'estimation statistique n'est possible que si l'échantillon est aléatoire et s'il est suffisamment important. L'estimation est interdite si l'échantillon est trop petit : n<30. Outre les difficultés pratiques du tirage aléatoire, la détermination de la taille de l'échantillon nécessite le choix d'un bon compromis entre la précision attendue et le coût de collecte des données.

Echantillons homogènes

On est dans cette situation lorsque les analyses ne portent que sur l'échantillon total. On cherchera dans ce cas à sélectionner l'échantillon le plus grand possible compte tenu des moyens disponibles. Plus l'échantillon est grand, plus il faudra interroger de personnes supplémentaires pour gagner en précision. A partir d'un certain seuil, les gains en précision ne justifient plus le coût supplémentaire que cela implique.

Tout dépend en fait du type de décision à prendre et de la marge d'incertitude tolérable. Ainsi, on sera beaucoup plus exigeant pour évaluer l'audience d'un média en vue de fixer des tarifs publicitaires que pour une étude de satisfaction. Dans le premier cas, on s'orientera vers de grands échantillons (1000 à 2000 sondés), dans le second, on pourra se contenter d'échantillons plus modestes (200 à 300).

Echantillons hétérogènes

Cette situation correspond au cas où l'on souhaite établir des résultats sur des sous-ensembles de l'échantillon. Si l'échantillon est aléatoire, on obtiendra des effectifs très faibles pour les catégories peu représentées ; ce qui risque d'interdire toute estimation sur les strates correspondantes. Ainsi, pour analyser la strate d'une catégorie représentant 2% de la population totale, il faut un échantillon contenant au moins 30 personnes dans cette catégorie ; ce qui nécessite l'interrogation de 1500 personnes. Mais cet effectif ne se justifie pas pour étudier la population dans son ensemble. Il faudra donc trancher entre renoncer à analyser toutes les strates ou supporter le coût d'interrogation de 1500 personnes.

Une solution de compromis consiste à définir un échantillon stratifié dans lequel on alloue le budget disponible à chacune des strates. Si on dispose d'un budget de 500 personnes et si la population se compose de 5 strates, on interrogera aléatoirement 100 personnes de chaque catégorie. On est ainsi assuré d'avoir une précision convenable pour l'analyse de chaque strate. Mais on ne pourra rien tirer de l'analyse de l'échantillon total dans lequel certaines strates seront sur-représentées et d'autres sous-représentées.

Redressement d'échantillon

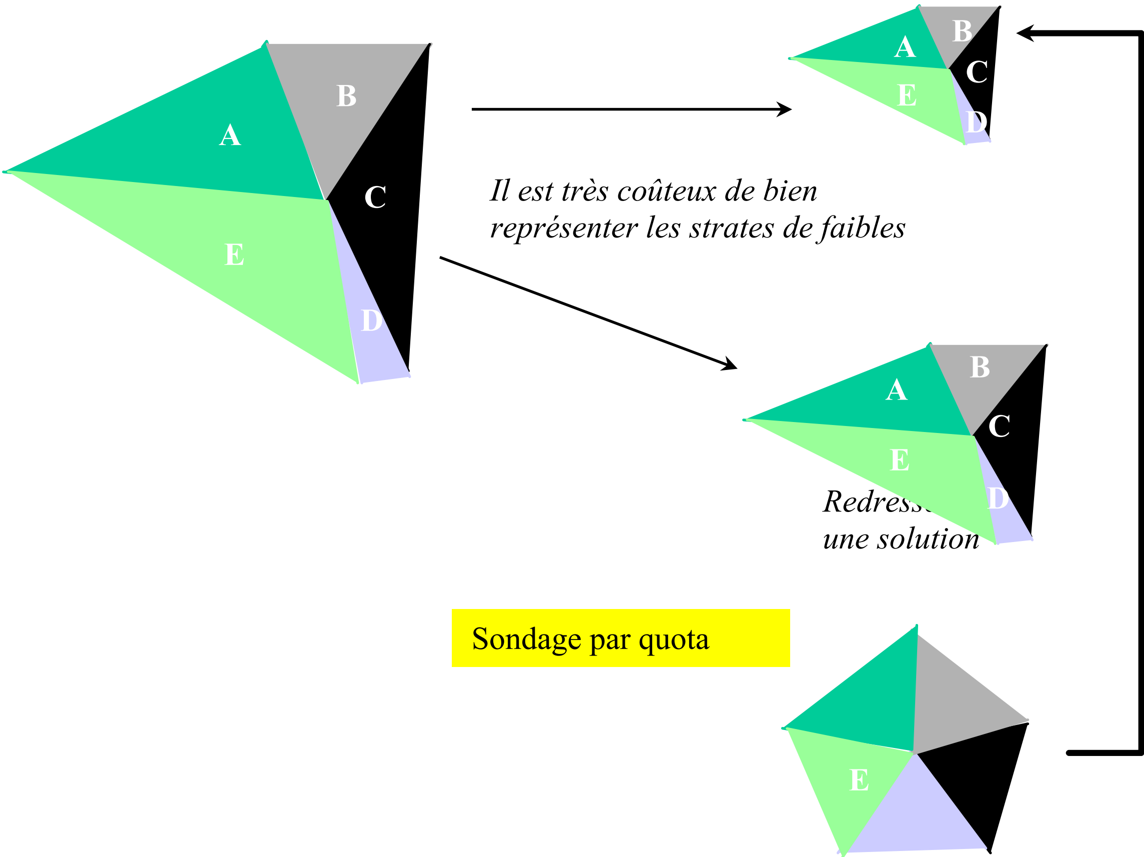
Dans le cas précédent, on redressera l'échantillon total pour composer un nouvel échantillon dans lequel chaque catégorie sera représentée à proportion de son poids dans la population totale.

Procéder ainsi conduit, au niveau de l'analyse, à travailler sur les échantillons propres à chaque strate et sur un échantillon redressé pour analyser la population totale.

Choisir la taille de l'échantillon

Taille de l'échantillon	**Taux d'erreur**	**Intervalle de confiance pour p=50%**
200	6,93%	[43,07% , 56,93%]
300	5,65%	[44,34% , 55,65%]
2000	2,19%	[47,81% , 52,19%]
3000	1,79%	[48,21% , 51,79%]

Sondage aléatoire



24. Les formes de l'analyse de données

Les méthodes de l'analyse de données sont multiples et répondent à des objectifs variés :

- dépouillement visant à restituer les réponses de manière individuelle ou synthétique ;

- transformation des données originales par recodification ou calcul ;

- analyses statistiques visant à décrire, expliquer ou classifier.

Elles peuvent concerner une ou plusieurs variables, un seul individu, tous les individus de l'échantillon ou un sous-ensemble appelé strate.

Le niveau d'analyse

Les dépouillements peuvent se faire à différents niveaux :

- Au niveau de l'individu. On s'intéresse aux données de détail en vue d'entreprendre des actions individuelles. Les traitements se ramènent alors à restituer tout ou une partie de l'information propre à chaque observation. Les résultats sont produits sous forme de listes plus ou moins étendues et structurées selon le nombre d'individus décrits. Ces extractions peuvent être effectuées au stade de la saisie ou du dépouillement en produisant des listes.

- Au niveau de la population entière. On cherche à synthétiser les informations décrivant chaque individu pour traiter la population comme un tout. Ceci revient à gommer les variations individuelles pour décrire chaque variable par un ou plusieurs indicateurs synthétiques : valeur moyenne pour les variables numériques, fréquence relative des modalités d'une variable nominale.

- Au niveau d'un sous-ensemble de la population. On analyse des catégories particulières d'individus pour tenir compte de l'hétérogénéité de la population. En définissant les conditions auxquelles doit répondre le sous ensemble étudié, on construit la strate qui se substitue alors à l'examen de l'échantillon total.

De l'univarié au multivarié

On peut distinguer 3 grandes problématiques auxquelles répondent 3 méthodes d'analyse :

- L'analyse univariée : on décrit la population en examinant une variable à la fois. C'est la manière la plus simple de restituer l'information et de faire de l'estimation statistique.

- L'analyse bivariée : on s'intéresse aux relations existant entre 2 variables à des fins d'explication et/ou de prédiction. Cette approche nécessite la formulation d'une hypothèse que la statistique permettra d'infirmer ou de confirmer.

- L'analyse multivariée : on aborde la complexité résultant de la multiplicité des variables.

Dans une approche descriptive, on cherche à réduire le nombre de variables (analyses factorielles) ou à agréger les individus en catégories homogènes (typologies).

Dans une approche explicative, on cherche à intégrer la pluralité des causes et des effets d'interaction (régression multiple – manova -).

La recodification

Les données disponibles ne sont pas toujours dans un format adéquat. La recodification permet de générer de nouvelles variables mieux appropriées à l'analyse. On peut distinguer :

La recodification d'une seule variable : mettre en classes une variable numérique, agréger les modalités d'une variable nominale, recoder une variable texte en fonction de son contenu.

La recodification de plusieurs variables : calculer un score à partir de plusieurs variables nominales, calculer une nouvelle variable avec une formule faisant intervenir plusieurs variables, créer de nouvelles variables calculées en application d'une analyse multivariée (scores factoriels, classifications...).

Les niveaux d'analyses

Analyser la population comme un tout : synthétiser

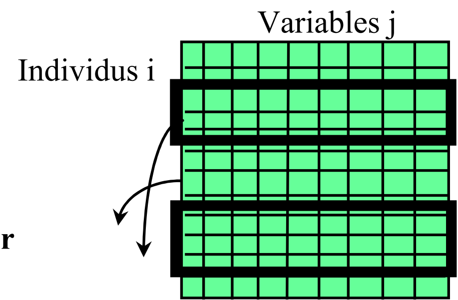
*On s'intéresse aux **variables***

Analyser les données individuelles: détailler

*On s'intéresse aux **individus***

Analyser les sous ensembles de la population: segmenter

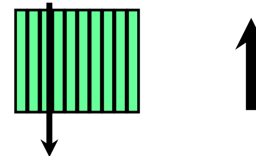
*On s'intéresse aux **strates***



Les problématiques

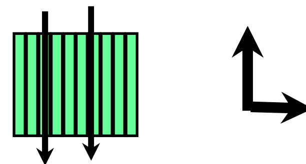
Analyse univariée

***Décrire** une variable à la fois*



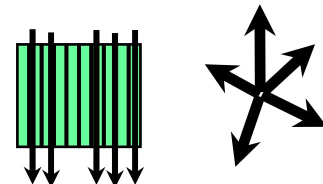
Analyse bivariée

*Mettre en relations 2 variables, **expliquer***



Analyse multivariée

*Analyser simultanément plusieurs variables, **synthétiser***



25. L'analyse univariée

L'analyse univariée consiste à donner une description synthétique de l'ensemble des individus observés ou d'un de ses sous-ensembles. La manière d'effectuer ces descriptions dépend de la nature statistique des variables en jeu. Dans le cas d'étude par sondage, on peut également se livrer à l'estimation des propriétés de la population mère.

Décrire

On analyse chaque variable pour parvenir à un énoncé synthétique du type :

" Il y a ...." : Il y a tant d'hommes, tant de femmes qui connaissent notre produit, il y a chez les acheteurs tant d'ouvriers...

" .... est (sont).... " : Le revenu moyen des ménages acheteurs est de.., la qualité est le premier critère de choix...

"...........fait (font)....." X% de l'échantillon fait confiance à......

Ces énoncés spécifient l'état d'une variable par un effectif (tant d'hommes), un pourcentage (X% de la population fait) une grandeur moyenne (le revenu moyen des ménages acheteurs est...) ou un rang (le premier critère de choix est la qualité). L'importance du phénomène considéré se trouve ainsi chiffrée. L'analyste peut, sur cette base, exercer son jugement pour décider. Si la précision le permet, il pourra effectuer des calculs et des prévisions.

Dans le cas d'une variable texte, la description consiste à restituer le texte sous forme de liste ou lexique indiquant la fréquence des termes utilisés. Dans tous les cas, il faudra lire et interpréter ces éléments dont le sens n'est pas, comme pour les autres variables, fixé a priori.

Les indicateurs selon le type de variable

Selon le type de variable, la nature de l'information recueillie n'est pas de même nature et n'autorise pas les mêmes traitements. A chaque cas correspondent des indicateurs de résultats différents.

Extrapoler à la population totale

Si l'échantillon répond aux conditions du sondage aléatoire (tous les individus de la population ont exactement la même chance d'être sélectionnés) et s'il comporte au moins 30 individus, on peut estimer à partir des indicateurs calculés sur l'échantillon, les propriétés de la population totale. La valeur de l'écart-type (numérique) et celle de l'intervalle de confiance (nominale) donnent les fourchettes de l'estimation.

Attention, dans le cas d'une strate, c'est l'effectif de la strate qu'il faut considérer et non pas celui de l'échantillon.

Le type de variable détermine les possibilités d'analyse

Type de variable	Nature de l'information	Opérations possibles	Indicateurs
numérique	nombre	calcul : +, -, x, : >, <, =, #	moyennes, écarts-types, effectifs, pourcentages
échelle	rang, ordre	comparaison : >, <, =, #	effectifs, pourcentages, sous certaines hypothèses moyennes et écarts-types
nominale	code : suite de caractères affectés d'une signification	identification : =, #	effectifs, pourcentages
texte	forme graphique : suite de caractères sans signification a priori	interprétation : ?	recodification, effectifs ou pourcentages de forme graphiques

26. L'analyse bivariée

L'analyse bivariée commence par la formulation d'une hypothèse orientée par la signification des variables et se poursuit par la mise en œuvre d'une méthode résultant de la nature des variables.

La sémantique : explication et relations causales

L'analyse bivariée conduit à formuler un énoncé de type Si..V_1.. Alors $V_{2...}$ postulant l'hypothèse d'une relation causale entre 2 variables. A ce stade, c'est la sémantique qui guide l'analyse : la signification des variables conduit à formuler une théorie justifiant la relation et son sens. La statistique confirme ou invalide son existence mais seule la théorie en donne le sens. Ainsi, on peut établir statistiquement un lien entre le niveau d'éducation et le revenu. Mais c'est en fonction d'une théorie qu'on interprétera cette relation pour dire que le revenu conditionne l'éducation (théorie du coût) ou que l'éducation détermine le revenu (théorie du rendement).

L'analyse bivariée commence donc par la formulation des hypothèses que la statistique permettra de tester. Le modèle suivant peut orienter la réflexion : parmi toutes les relations envisageables, 3 renvoient à des théories très générales du comportement humain :

1 - Le déterminisme sociologique : l'action obéit aux habitudes et aux contraintes.

2 - La décision et la rationalité : l'action résulte des choix et des calculs.

3 - Les psychologismes : l'action est modalité d'expression.

On peut bien sûr faire l'économie de la réflexion préalable et essayer toutes les relations envisageables. Elles peuvent être très nombreuses et perdre l'analyste dans une quête aveugle. D'autre part, le fait de constater une relation statistique ne suffit pas à établir une connaissance argumentée.

Avant de commencer toute analyse bivariée, il convient d'établir une stratégie de recherche, en mobilisant les expériences, intuitions, croyances, théories, toutes connaissances préalables que l'on confrontera aux informations contenues dans la base de données. On déterminera ainsi quelles variables mettre en relation. La statistique conduira à rejeter l'hypothèse si on ne peut pas montrer que la relation recherchée existe. Dans le cas inverse, la qualité de l'interprétation du fait statistique ne dépendra que de celle de la théorie utilisée.

La statistique : mettre en œuvre la méthode adaptée

La méthode à mettre en œuvre pour tester l'existence d'une relation dépend de la nature des variables en présence. Le tableau suivant identifie les 3 cas possibles selon la nature des variables (nominales ou numériques).

Pour les variables échelles, on se ramène à l'un des cas précédents selon qu'on les considère comme des variables numériques ou comme des variables nominales.

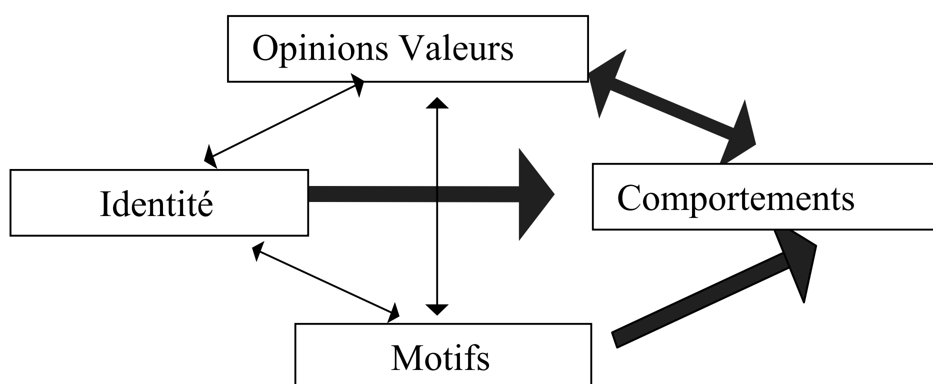
Le cas des variables textuelles peut également être ramené à l'un des 3 cas :

- En faisant une analyse de contenu par recodification de la variable, on est ramené au cas du croisement avec une variable nominale.

- En utilisant les méthodes de l'analyse lexicale : construction d'un tableau lexical qui décompte le nombre d'occurrences des mots de la variable texte, pour les modalités d'une variable nominale. On est ramené au traitement d'un tableau de contingence analogue au tri croisé,

- création de nouvelles variables d'origine lexicale : nominales (fermées sur les mots du lexique) ou numériques (longueur, banalité, intensité...) susceptibles d'être mises en relation entre elles ou avec les variables de contexte) ; ce qui ramène à l'un des cas précédents.

La sémantique : la signification des variables guide la formulation des hypothèses



La statistique : la nature des variables détermine la méthode

	V2 : NOMINALE	V2 : NUMERIQUE
V1 : NOMINALE	**1 Tris croisés** Recherche de correspondances par le calcul des effectifs relatif aux croisements des modalités de chaque variable (tableaux de contingence). Test du chi2 : analyse des correspondances. Séquences : ***Analyser / Tableaux croisés*** ***Analyser / Tableaux multiples***	**3 Analyse de la variance** Recherche des variations de V2 selon les modalités de V1 par le calcul des moyennes et écart-types de V2 sur les strates définies par les modalités de V1. Test de comparaison de moyennes, analyse de la variance. Séquence : ***Approfondir / Tableau de moyennes croisées***
V1 : NUMERIQUE	**3 Analyse de la variance** Recherche des variations de V1 selon les modalités de V2 par le calcul des moyennes et écart-types de V1 sur les strates définies par les modalités de V2. Test de comparaison de moyennes, analyse de la variance. Séquence : ***Approfondir / Tableau de moyennes croisées***	**2 Corrélation** Recherche d'une relation graphique ou algébrique entre les valeurs des 2 variables (droite et équation de régression). Test de corrélation : analyse de la relation. Séquence : ***Analyser / Corrélation et nuage de points***

27. L'analyse bivariée : Chi2 et AFC

Afin de déterminer s'il existe un lien entre 2 variables nominales, on compare les effectifs du tableau à ceux qu'on aurait obtenus si les effectifs étaient répartis proportionnellement. Si tel est le cas, quelle que soit la modalité considérée d'une des variables, la répartition des modalités de l'autre reste la même. Il y a alors indépendance entre les 2 variables.

Le calcul du Chi^2

Le test du Chi² consiste à déterminer si le tableau étudié correspond à cette hypothèse. S'il s'en écarte beaucoup, on présume l'existence d'un lien d'autant plus significatif que l'écart est grand. Le Chi² est la mesure de cet écart.

A partir d'un tableau de contingence à L lignes Li et C colonnes Cj, on détermine d'abord, pour chaque case, l'effectif théorique. C'est le produit du total de sa ligne (Li,.) par le total de sa colonne (C.,j) divisé par le total général (n) soit (Li,.*C.,j/n). On calcule ensuite la somme des carrés des écarts entre effectif réel et effectif théorique. Plus l'écart est grand, plus le Chi² est élevé, et plus on a de chances d'être en présence d'un lien significatif.

Cette appréciation dépend bien sûr de la dimension du tableau, c'est-à-dire du nombre de degrés de liberté : ddl = (L-1)* (C-1). Plus il y a de cases, plus la somme risque d'être élevée. Si la valeur du Chi² permet d'indiquer l'existence d'un lien, il faut, pour le qualifier, examiner comment celui-ci est composé.

Contribution au Chi² et interprétation des correspondances

Sur quelles cases observe-t-on les écarts les plus importants ? Sur quelles autres les effectifs sont sans surprise ?

On examine pour cela les contributions de chaque case à la somme du Chi², elles mettent en évidence l'importance de l'excès ou du déficit observable dans chaque cellule. Les cases contribuant le plus fortement sont encadrées de bleu ou de rouge selon que l'effectif réel excède ou est inférieur à l'effectif théorique.

Ainsi, c'est l'examen des contributions au Chi² et des correspondances qu'elles révèlent qui permet véritablement de qualifier la relation.

Construction et lecture d'une carte d'AFC

On peut donner une représentation plus visuelle des écarts à l'indépendance par la technique de l'analyse factorielle des correspondances. Elle conduit à tracer une carte qui dispose les modalités des 2 variables en fonction des écarts à la situation d'indépendance. Par défaut, chaque modalité est représentée par un pavé de surface proportionnelle à son effectif. Leurs positions les unes par rapport aux autres s'interprètent ainsi :

- 2 modalités lignes et colonnes seront d'autant plus proches que les effectifs du tableau sont en excès par rapport à l'indépendance : attraction.

- Les modalités lignes et colonnes seront d'autant plus éloignées que les effectifs du tableau sont en déficit par rapport à l'indépendance : répulsion.

Les modalités lignes ou colonnes situées à la périphérie de la carte signalent des profils originaux. Au contraire, une position centrale interdit tout commentaire (profils sans originalité ou point mal représenté dans le système d'axes de la carte).

Le bien fondé de ces interprétations dépend de :

- L'intensité du lien entre les 2 variables, mesuré par le Chi².
- La quantité d'informations restituée par la carte, indiquée par le pourcentage de variance expliquée (ou d'écart à l'indépendance) par les axes. La qualité de la représentation est d'autant meilleure que ces pourcentages sont élevés.
- L'interprétation des axes à partir des oppositions qu'ils mettent en évidence doit tenir compte du pourcentage de variance restituée. S'il est faible, il faut se garder d'insister sur des phénomènes qui ne représentent qu'une petite partie des caractéristiques du tableau.

Hypothèse : Le point de vue privilégié à l'achat a une influence sur la marque choisie. (Référence théorique : la décision rationnelle : Si Motifs ---> Comportement)

MARQUE x CRITERES

- Quelle est la marque de la voiture que vous utilisez ?
- Quels sont pour vous dans la liste suivante les trois principaux points à prendre en considération pour l'achat d'une automobile ?

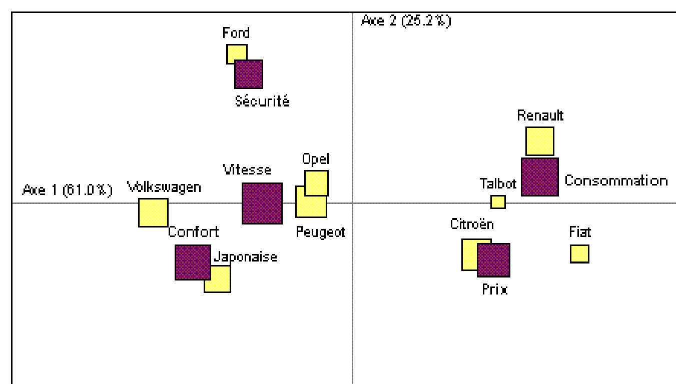
Tableau de contingence

CRITERES / **MARQUE**	Vitesse	Confort	Sécurité	Consom mation	Prix	**TOTAL**
Renault	22%	0%	17%	100%	17%	**100%**
Peugeot	64%	64%	36%	50%	29%	**100%**
Citroën	44%	17%	0%	56%	72%	**100%**
Talbot	25%	0%	25%	50%	75%	**100%**
Ford	50%	10%	80%	10%	10%	**100%**
Fiat	0%	17%	0%	83%	83%	**100%**
Volkswager	59%	55%	23%	0%	0%	**100%**
Opel	64%	36%	36%	36%	27%	**100%**
Japonaise	44%	61%	6%	6%	17%	**100%**
TOTAL	**45%**	**27%**	**22%**	**49%**	**35%**	**100%**

Test du chi2
Les données ne contredisent pas l'hypothèse

La dépendance est très significative. chi2 = 121,09, ddl = 32, 1-p = >99,99%.
Les cases encadrées en bleu (rose) sont celles pour lesquelles l'effectif théorique est nettement supérieur (inférieur) à l'effectif réel.
Attention, 26 cases ont un effectif théorique inférieur à 5, les règles du chi2 ne sont pas réellement applicables..

Analyse factorielle des correspondances



28. L'analyse bivariée : corrélation

Lorsque les variables sont numériques, on dispose d'une information très riche autorisant la recherche d'une formule mathématique pour qualifier la relation.

Nuage de points et droite de régression

La recherche d'une relation entre 2 variables numériques ***x*** et ***y*** peut se faire de 2 manières différentes :

D'une manière graphique, en représentant chaque observation par ses coordonnées **x** et **y** selon 2 axes. On obtient alors un nuage de points plus ou moins bien alignés.

D'une manière algébrique, en recherchant l'existence d'une relation linéaire entre ces 2 variables **y** = a**x** + b. On obtient des valeurs calculées **y** plus ou moins proches des valeurs observées.

Selon que la forme du nuage est plus ou moins proche d'une droite (la droite de régression), ou que les valeurs calculées à partir de l'équation (de régression) sont plus ou moins proches des observations réelles, on dira que la corrélation entre les 2 variables est bonne ou mauvaise. Le coefficient de corrélation mesure la qualité de l'ajustement entre les valeurs **y** et **x** réelles et le modèle de la relation représenté par l'équation **y** = a**x** + b ou par la droite correspondante. En référence à une interprétation causale du modèle, **y** est appelée variable à expliquer et **x** variable explicative.

Coefficient de corrélation

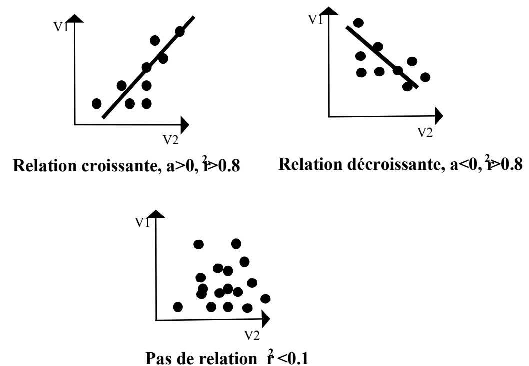
Le coefficient de corrélation (*r*) établit la qualité de l'ajustement entre deux variables V1 et V2. Sa valeur absolue (ou r^2) est comprise entre 0 et 1.

Elle est égale à 1 si l'ajustement est parfait : il existe une fonction V1=*a*xV2+*b* dont le résultat donne toujours exactement la valeur observée de V1 : on peut alors dire que V1 dépend exactement de V2.

Elle est égale à 0 si quelle que soit la fonction considérée, la valeur calculée de V1 est également éloignée de sa valeur observée. V1 est indépendante de V2.

L'usage est de considérer qu'à partir d'un coefficient de corrélation de valeur absolue supérieure à 0.8, il existe une bonne relation entre les 2 variables.

Le coefficient de corrélation est également affecté d'un signe. Il indique le sens de la relation. Elle est croissante si le signe est positif, décroissante sinon. Ce signe est aussi celui de *a*, le coefficient de régression dans l'équation V1=axV2+b.



Relation croissante, a>0, r^2>0.8 **Relation décroissante, a<0, r^2>0.8**

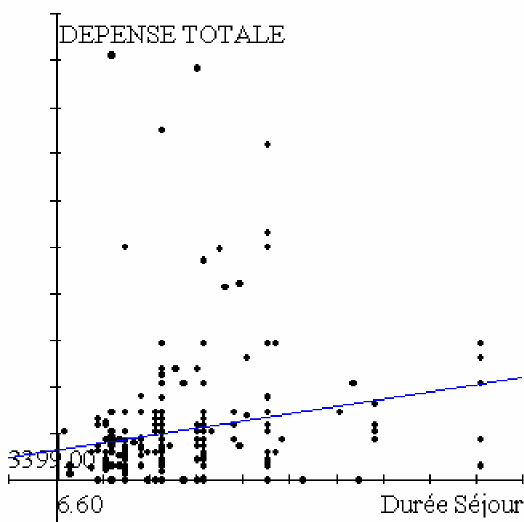
Pas de relation r^2<0.1

Dans le cas d'une régression multiple, on cherche à établir une relation du type V1=*a*xV2+*b*xV3+*c*xV4. On calcule alors un coefficient de régression multiple. Il indique également la qualité de l'ajustement effectué par le modèle et s'interprète comme un coefficient de corrélation simple.

Modèle linéaire et non linéaire

Le calcul du coefficient de corrélation se fait par rapport au modèle linéaire de l'équation du premier degré à laquelle correspond la droite de régression. On peut tester la relation par référence à un modèle non linéaire.

Il faudra pour cela transformer la valeur de la variable expliquée en la recalculant par rapport au modèle à tester. Par exemple, pour tester un modèle logarithmique, on calculera une nouvelle variable égale au log de la variable explicative : log(x) pour ensuite tester l'équation y=a.logx+b.

Hypothèse : Plus le séjour est long, plus la dépense totale est élevée (Théorie comptable élémentaire de la proportionnalité des coûts).	Corrélation. DEPENSE TOTALE /.DUREE SEJOUR
Nuage de points et droite de régression	 *209 individus ne sont pas pris en compte (non-réponse à au moins un des critères).* *Chaque individu est représenté par un point.*
Equation de la droite de régression	*Le graphique montre les 349 points de coordonnées DUREE SEJOUR; DEPENSE TOTALE* Equation de la droite de régression : DEPENSE TOTALE = 81.05 * Durée Séjour + 2210.08
Test de corrélation Les données contredisent l'hypothèse	*Coefficient de corrélation : 0.198 (Durée Séjour explique 3 % de la variance de DEPENSE TOTALE)* *Ecart-type du coefficient : 21.588*

29. L'analyse bivariée : analyse de la variance

L'analyse de la variance s'applique au cas où les variables sont de natures différentes : l'une est nominale et définit des catégories d'individus, l'autre, numérique, permet de mesurer une propriété de ces individus. On pourra ainsi par exemple savoir si, selon le sexe, le revenu varie significativement.

Tableau de valeurs moyennes

La construction d'un tableau de valeurs moyennes, établit, pour les modalités d'une variable nominale, les valeurs moyennes d'une ou plusieurs variables numériques. On peut ainsi comparer ces valeurs entre elles et conclure à l'existence d'une relation si les variations de la moyenne mettent en évidence des différences significatives sur l'ensemble des modalités. On utilise alors le test de l'analyse de la variance.

Analyse de la variance

Son but est d'établir si, au regard des valeurs de la variable numérique, les groupes d'individus correspondant aux modalités de la variable nominale sont significativement différents les uns des autres. Elle met en œuvre les principes suivants :

Pour chaque modalité de la variable nominale, la moyenne de la variable numérique dissimule une dispersion autour de cette moyenne. La variance (le carré de l'écart-type) mesure cette dispersion appelée **variance interne**.

D'une modalité à l'autre, la moyenne varie et révèle une hétérogénéité, plus ou moins grande, mesurée par un autre calcul de variance : **la variance entre modalités.**

Le test d'analyse de variance porte sur le rapport entre la variance entre modalités et la moyenne des variances internes.

Il a pour but de vérifier si l'hétérogénéité entre modalités est plus grande que l'hétérogénéité à l'intérieur des modalités. Si ce rapport noté F est suffisamment élevé, on dit que la variable numérique discrimine les modalités de la variable nominale.

Ce jugement global s'applique à la répartition de la population totale en catégories définies par chacune des modalités. On peut le compléter en comparant les valeurs moyennes de chaque modalité à la moyenne établie sur l'ensemble de la population.

Critères discriminants et valeurs significatives

Les tests d'analyse de la variance et de comparaison de moyennes apportent une information très utile permettant d'identifier les variables pour lesquelles les catégories de la variable nominale font apparaître des différences discriminantes :

Les variables dont le nom est encadré de bleu discriminent les modalités de la variable nominale. Pour ces variables, le test de Fisher est significatif (par défaut au risque de 5 %).

Les cellules encadrées dans le tableau signalent une moyenne significativement différente de la moyenne sur l'ensemble de l'échantillon. Le test de comparaison de moyennes est significatif (par défaut au risque de 5 %).

Représentation graphique de la dispersion

L'analyse d'un tableau de moyennes est facilitée par la représentation graphique de la dispersion des variables numériques selon les modalités de la nominale.

Analyse de la variance à plusieurs facteurs

On peut mener une analyse de la variance en considérant 2 variables nominales et une numérique.

On cherche alors, par la comparaison de tous les cas définis par les nominales les influences directes et croisées qu'elles peuvent avoir sur la valeur de la numérique. Cette méthode dite analyse de la variance multiple (Manova) fait partie des méthodes multivariées. Elle n'est disponible que dans Plus[2].

Hypothèse : Le mode d'hébergement a une influence sur les dépenses et la durée du séjour (référence au mode de production et élasticité prix).

Tableau de moyennes de.MODE HEBERGEMENT

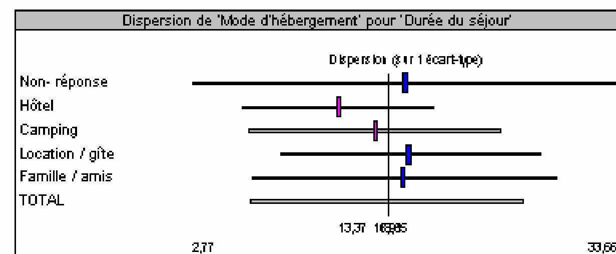
Critères évalués : DUREE DU SEJOUR, DEPENSE TOTALE

Tableau des valeurs moyennes pour Durée du séjour Dépense totale et Dépense d'habillement.

Mode d'hébergement	Durée du séjour	Dépense totale
Hôtel	13,37	5007,26
Camping	16,00	2536,15
Location / gîte	18,55	4643,26
Famille / amis	18,16	2803,65
TOTAL	16,91	3588,18

Les valeurs du tableau sont les moyennes calculées sans tenir compte des non-réponses.
Les noms des critères discriminants sont encadrés.
Les nombres encadrés correspondent à des moyennes par catégorie significativement différentes (test t) de l'ensemble de l'échantillon (au risque de 5%).

Graphique de dispersion



Test d'analyse de la variance : l'hypothèse est confirmée : le mode d'hébergement a une incidence sur la durée du séjour et sur la dépense.

Résultats du test de Fisher :
Durée du séjour : V_inter = 505,59, V_intra = 94,96, F = 5,32, 1-p = 99,96%
Dépense totale : V_inter = 101930198,55, V_intra = 14443866,90, F = 7,06, 1-p = >99,99

30. Analyse multivariée : objectifs et méthodes

L'analyse multidimensionnelle des données consiste à analyser simultanément plus de deux variables à la fois dans un but de synthèse ou d'analyse.

Synthétiser

Il s'agit de résumer la masse des informations concernant un grand nombre d'individus décrits par de nombreuses variables. On cherche à décrire les données par une expression plus économique afin d'en rendre compte plus simplement. Deux voies sont envisageables :

La classification ou la typologie : elle consiste à regrouper les individus de la base de données en classes (ou types) d'individus selon les similarités qui permettent de les rassembler. Par exemple : décrire une population en identifiant différentes classes de styles de vie.

La réduction des dimensions d'analyse : elle consiste à réduire un grand nombre de variables à quelques dimensions communes. Par exemple, ramener la variété des opinions exprimées par les réponses à vingt questions différentes, à 2 dimensions opposant d'une part, le sentiment à la raison, et d'autre part, la tradition au changement.

Analyser

On cherche les influences de plusieurs variables entre elles pour mettre en évidence comment celles-ci se déterminent les unes et les autres.

On cherche à expliquer en construisant des modèles permettant d'expliquer une variable par les variations de plusieurs autres, en cherchant ainsi des liens de causalité plus complexes que la simple relation entre deux variables.

Les méthodes

Les méthodes disponibles sont nombreuses.

Certaines sont directes. Elles consistent à :

- combiner entre elles plusieurs variables pour calculer une nouvelle variable qui les résume (une somme, une moyenne, un score...).

- marquer des individus selon leur appartenance à telle strate ou selon qu'ils répondent à tel profil et constituer ainsi des groupes a priori.

D'autres sont indirectes. Elles reposent sur une analyse préalable de la structure des données orientant la synthèse en fonction des propriétés révélées. On distingue ces méthodes selon le type de variables auxquelles elles s'appliquent :

- Pour les variables quantitatives, il s'agit de :

La corrélation multiple : elle permet de construire la matrice de corrélation qui met en évidence l'ensemble des relations existant entre plusieurs variables.

L'analyse en composantes principales : à partir des propriétés de la matrice de corrélation, elle permet de substituer aux variables de départ un plus petit nombre de dimensions qui les résument au mieux. Dans cet espace réduit, on peut définir une typologie en regroupant les individus en classes homogènes.

La régression multiple : à partir des propriétés de la matrice de corrélation, elle permet de bâtir un modèle explicatif donnant la meilleure expression de l'une des variables - variable à expliquer - en fonction des autres - variables explicatives -.

La classification automatique : cette méthode consiste à répartir les individus d'une population en un nombre de classes déterminé a priori. En fonction de la structure des données, la méthode consiste à améliorer une partition initiale des individus.

- Pour les variables qualitatives, il s'agit de :

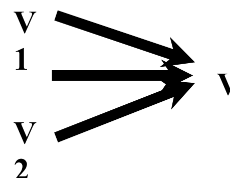
L'analyse factorielle des correspondances multiples : à partir d'une analyse des distances entre individus, définie par rapport à leur description sur un grand nombre de variables qualitatives, on détermine un sous-ensemble de dimensions, des facteurs conservant au mieux les distances de départ. Dans cet espace réduit, on peut alors définir une typologie en regroupant les individus en classes homogènes.

La classification automatique appliquée à des variables qualitatives.

Toutes ces méthodes sont accessibles par le menu ***Approfondir***. On peut pour mémoire y rajouter des méthodes de même nature, ne portant pas sur les individus mais sur des tableaux d'effectifs ou des tableaux de moyennes. Le tableau ci-dessous résume les méthodes disponibles.

Explique

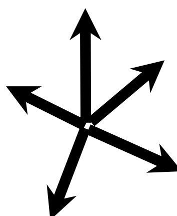
Etablir un modèle des relations entre variables. Il exprime ces relations sous forme de fonction mathématique.

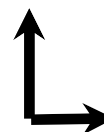


Régression multiple : n variables numériques
Manova : 2 variables nominales et 1

Synthétise

Réduire le nombre des dimensions d'analyse : cartes factorielles, scores factoriels.

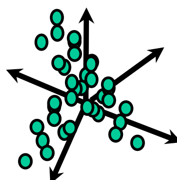


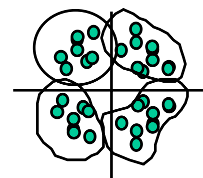


Analyse factorielle multiple (**AFCM**) : n variables nominales
Analyse en composantes principales (**ACP**) : n variables numériques
Analyse factorielle des correspondances (**AFC**) : 2 nominales
Analyse en composantes principales
d'un tableau de moyennes (**ACP**) : 1 nominale n numériques

Classifie

Regrouper les individus en classes homogènes. Classification automatique, typologie à vue.





Classification automatique : n numériques
Typologie AFCM : n nominales
Typologie ACP : n numériques

31. L'analyse factorielle : les principes

Les méthodes d'analyse factorielle peuvent porter sur des données de dimensions modestes, tableaux croisés, tableaux de moyennes… ou sur des ensembles de grande dimension lorsqu'en ligne, on trouve tous les individus de la population. Dans ce dernier cas, la recherche des facteurs peut se prolonger par la construction d'une typologie.

AFCM et ACP

Les deux méthodes présentées ici portent sur ce type de données comportant autant de lignes que l'échantillon ou la strate compte d'individus.

Quant aux colonnes, ce sont :

- soit des variables numériques ou critères (questions ouvertes numériques ou échelles), auquel cas, il s'agit d'un tableau de valeurs (chaque case est la valeur donnée par l'individu en réponse à la question posée) et l'analyse est une **Analyse en Composantes principales** ou ACP sur individus.

- soit des variables qualitatives (questions fermées uniques ou multiples), auquel cas, elles représentent les modalités. Le tableau est alors un tableau binaire ne comportant que des 1 ou des 0, selon que l'individu a cité ou non la modalité en question. Un tel tableau binaire peut être assimilé à un tableau d'effectifs et peut être soumis comme tel à l'**Analyse factorielle des correspondances multiples** ou AFC sur individus.

Bien qu'elles concernent des variables de natures différentes, ces deux méthodes mettent en œuvre les mêmes principes et la même démarche. Nous les décrivons d'une manière commune dans ce qui suit, pour illustrer plus en détail ensuite l'***Analyse en composantes principales*** et situer enfin, par différence, les spécificités de l'***Analyse factorielle multiple***.

Les données individus / variables

Le tableau de *i* lignes, représentant les individus, et de *c* colonnes, représentant les dimensions, correspond à l'univers à analyser. Il forme un hyper-espace de *c* dimensions dans lequel se situent les *i* individus.

Les techniques qui nous intéressent ont pour but de ramener cet espace à de plus modestes dimensions. L'idéal est de le réduire aux deux dimensions d'un plan. Cette configuration permet en effet de visualiser les individus et de mettre en évidence ce qui les distingue. Ces deux dimensions, appelées facteurs, résument l'ensemble des variables de départ et peuvent conduire à une interprétation plus synthétique de la structure des données.

A partir de ces mêmes informations, on peut regrouper les individus en classes de proximité pour définir ainsi une typologie. Cette démarche est illustrée par le schéma suivant.

Trouver les facteurs pour réduire le nombre des variables d'analyse

La réduction de l'espace de départ se fait par la recherche des facteurs résumant l'ensemble des données.

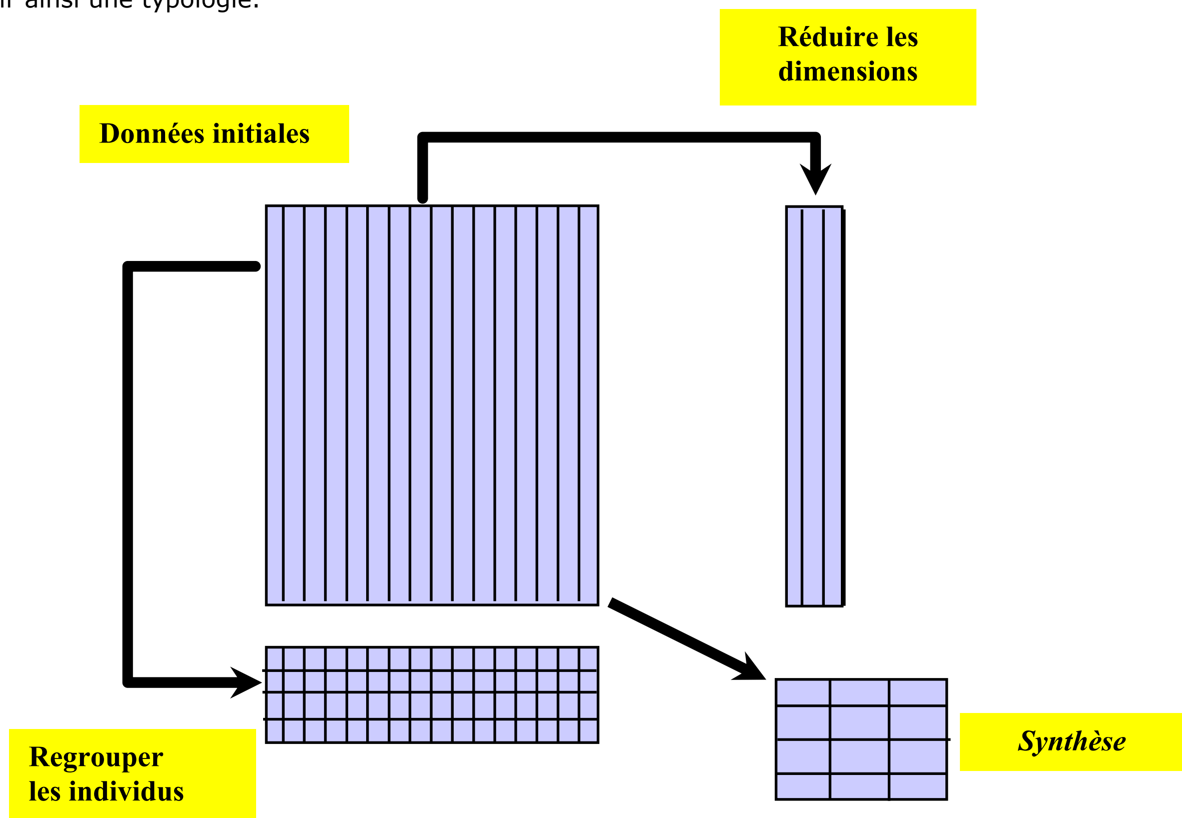
Pour l'***ACP*** - cas des données quantitatives -, on cherche les combinaisons linéaires des variables qui conservent le mieux la structure des données. On s'appuie sur l'analyse de la matrice de corrélation. Pour l'***AFM*** – cas des données qualitatives -, on procède par une analyse de la distance du Chi² en vue de déterminer les facteurs qui restituent au mieux les écarts entre individus. On s'appuie sur le tableau des effectifs ou tableau de Burt.

Agréger les individus : construire une typologie

La détermination des facteurs résumant l'information de départ permet, par projection, de représenter les individus dans un plan. Les individus apparaissent ainsi sous la forme d'un nuage de points mettant en évidence différentes zones. On peut regrouper les individus en catégories selon leur disposition sur la carte et définir ainsi une typologie.

Scores factoriels et typologie

L'enregistrement des coordonnées des individus par rapport aux axes factoriels (score factoriels) et leur appartenance à l'une des catégories de la typologie permet de résumer l'information de départ.



Variables numériques : **Analyse en composantes principales**

Variables nominales : **Analyse factorielle multiple**

On enrichit la base de départ en créant de nouvelles variables correspondant aux :

Scores factoriels
Classes typologiques

32. L'analyse factorielle : interpréter les résultats

Qualité de la réduction par l’analyse factorielle

Les algorithmes utilisés consistent à rechercher les n premiers facteurs. Chacun d’eux restitue une partie de l’information du tableau de départ. Le premier facteur en restitue le plus et ainsi de suite, selon l’ordre décroissant de la quantité d’informations.

La qualité de l’analyse dépend ainsi de l’information restituée par les deux premiers facteurs. Plus celle-ci est élevée, meilleure sera l’analyse effectuée dans le plan qu’ils définissent. Lorsque celle-ci est trop faible, il peut être nécessaire de la compléter par l’analyse des plans suivants.

La quantité d’informations restituée par un facteur (ou axe factoriel) est indiquée par le pourcentage de variance expliquée par le facteur (AFC) ou la composante (ACP).

Interpréter les facteurs

L’interprétation des facteurs peut se faire par l’examen du plan factoriel ou du tableau des contributions. Les contributions indiquent dans quelle mesure les axes restituent l'information contenue dans les variables de départ. Visuellement, on peut raisonner comme si, par leur position éloignée du centre, les variables (ACP) ou les modalités (AFCM) « tiraient les axes factoriels en leur donnant leurs propriétés ». Au contraire, lorsqu’elles sont près du centre, elles n’ont pas ou peu d’influence.

Dans le cas de l'ACP, les coordonnées des variables sont égales au coefficient de corrélation de la variable avec chacun des axes. Pour l'AFCM, elles sont un indicateur de la contribution des modalités à chacun des axes.

Trouver une typologie pour grouper les individus en classes homogènes

Par projection, on peut représenter les individus dans le plan factoriel, mais selon leur position dans l’espace d’origine, ils seront plus ou moins bien représentés. S’ils sont éloignés du plan de projection, ils se projetteront au centre du plan et leur position fera illusion.

Dans le cas de l'ACP, les coordonnées des variables sont égales au coefficient de corrélation de la variable avec chacun des axes. Pour l'AFCM, elles sont un indicateur de la contribution des modalités à chacun des axes.

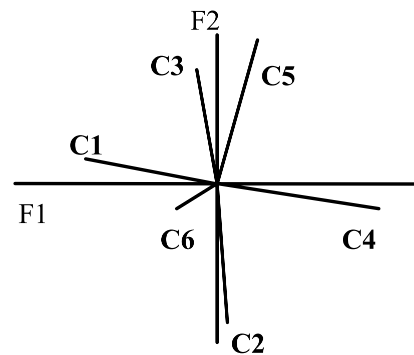
Ainsi dans la figure suivante, les individus C et A semblent proches alors qu’ils sont en fait éloignés. A est mal représenté car très éloigné du plan factoriel avec lequel il forme un angle presque droit.

La projection est déformante, il convient donc de procéder avec prudence, par exemple en ignorant les points mal projetés (ceux qui font un angle élevé avec le plan) et en concentrant l’analyse sur les groupes de points éloignés du centre.

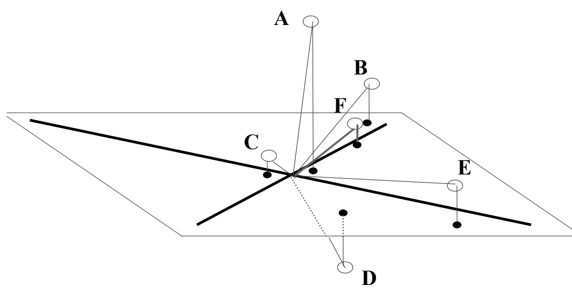
Ces précautions étant prises, il est alors possible de regrouper les individus, selon leur proximité dans le plan factoriel, en construisant ainsi une typologie à vue. Cette possibilité est une des originalités du Sphinx. Elle est accessible à partir du bouton ***Typologie*** figurant en regard des cartes d’analyses factorielles.

Interpréter les axes

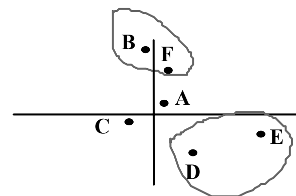
Le schéma ci-contre conduit à interpréter le facteur F1 comme la dimension sur laquelle s'opposent les critères C1 et C4, alors que sur le facteur F2, C3 et C5 s'opposent à C2.
Les variations du critère C6 sont mal représentées par ce plan factoriel.
L'examen du tableau des contributions associé à cet exemple confirmerait ces interprétations. Par exemple, on pourrait y vérifier que C6 ne contribue que pour un très faible pourcentage à chacun des axes.



Projection des individus dans le plan factoriel



Vue dans l'espace



Vue dans le plan factoriel

33. Corrélation et régression multiple

L'objectif de la régression multiple est de mettre en relation une variable, la variable expliquée avec plusieurs autres les variables explicatives, dans le but de définir l'équation qui les relie. On pourra ainsi, connaissant les variables expliquées, déterminer les valeurs de la variable explicative : V1=*a*xV2+*b*xV3+*c*xV4. On calcule alors un coefficient de régression multiple. Il indique la qualité de l'ajustement effectué par le modèle et s'interprète comme un coefficient de corrélation simple.

L'exemple ci-contre, tiré d'une étude de satisfaction, montre comment on peut expliquer la satisfaction globale en la rapportant à l'évaluation des différents attributs du service considéré. Les coefficients de régression partielle (a_i) donnent une indication sur la manière dont la perception de chaque attribut intervient dans la formation de la satisfaction globale.

Equation de régression linéaire multiple et paramètres d'ajustement

On calcule l'équation linéaire qui ajuste le mieux la variable expliquée par rapport aux variables explicatives. Les résultats sont communiqués sous la forme de l'équation de régression multiple.

La qualité de l'ajustement s'apprécie principalement à la valeur du coefficient de corrélation. Plus sa valeur absolue est élevée, plus faible est l'écart entre les valeurs calculées et observées (cet écart est aussi appelé résidu).

L'effet de chaque variable explicative dépend des coefficients de régression figurant dans l'équation. Plus celui-ci est grand, plus la variable explicative considérée influence la variable expliquée. Mais il faut également tenir compte de l'écart-type de chacun de ces coefficients. Plus il est élevé, moins nette est l'influence de la variable considérée.

Enfin, l'indicateur F est un autre moyen d'apprécier la qualité de l'estimation. Si sa valeur est supérieure au seuil d'une table de Fisher, l'estimation est considérée comme très significative (à 95%) ou peu significative (entre 80 et 95%), sinon, elle ne l'est pas du tout.

Les variations de F peuvent conduire à reconsidérer les variables intervenant dans le modèle. En effet, la suppression ou l'ajout de variables supplémentaires peut affecter la qualité de l'ajustement (coefficient de corrélation) mais dégrader celle de l'estimation. On observe alors une décroissance de F.

Procédure de régression pas à pas

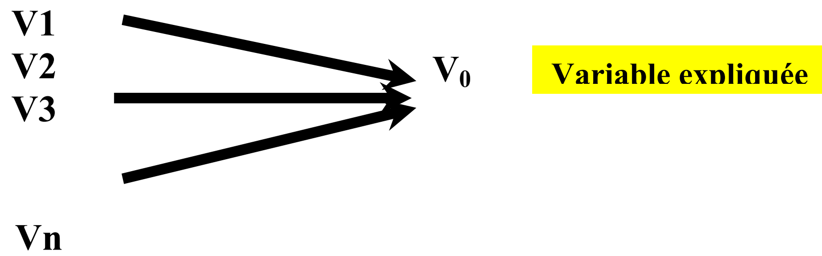
Les variables explicatives interviennent dans le calcul, dans l'ordre de leurs corrélations respectives avec la variable à expliquer. L'algorithme s'arrête quand la qualité de l'ajustement n'augmente plus de manière significative.

Indépendance des variables explicatives

L'examen de la matrice des coefficients de corrélation permet de vérifier que les variables explicatives intervenant dans l'équation sont indépendantes. On peut, en outre, tester la sensibilité de la régression en supprimant ou rajoutant des variables. On peut alors suivre l'évolution du coefficient de corrélation multiple.

Régression non linéaire

Il est tout à fait possible de tester des modèles non linéaires. Il faut pour cela au préalable transformer les variables intervenant dans le modèle en utilisant la fonction.



Variables explicatives

$$Vo = a_1 x V_1 + a_2 x V_2 + a_3 x V_3 \ldots + a_n x V_n$$

Note globale

a1, a2, a3… coefficients d'importance relative

Equation de régression multiple

note_globale = +0.970 * Restauration +0.891 * Confort -0.652 * Soins_qualité +0.639 * Informations -0.510 * Réconfort +0.390 * Hygiène -0.219 * Services +9.709

Matrice de corrélation

	note _glo	Rest aura	Conf ort	Hygi ène	Infor mati	Soin s_ra	Com péte	Disp onib	Réc onfo	Acc ueil	Serv ices	Soin s_q
note_globale	1,00											
Restauration	0,42	1,00										
Confort	0,35	0,17	1,00									
Hygiène	0,26	0,33	0,09	1,00								
Informations	0,23	0,12	0,05	0,12	1,00							
Soins_rapidité	0,15	0,22	0,16	0,19	0,17	1,00						
Compétence	0,15	0,22	0,25	0,08	0,08	0,17	1,00					
Disponibilité	0,14	0,26	0,17	0,22	0,22	0,31	0,15	1,00				
Réconfort	-0,00	0,15	0,18	0,29	0,12	0,18	0,14	0,20	1,00			
Accueil	-0,06	-0,04	-0,02	-0,03	0,02	-0,09	-0,00	-0,06	0,03	1,00		
Services	-0,06	0,04	-0,03	0,18	-0,00	0,03	-0,02	0,04	0,00	0,02	1,00	
Soins_qualité	-0,14	0,15	0,10	-0,04	0,07	0,09	0,08	0,08	0,01	-0,04	0,04	1,00

34. Classification automatique

La classification automatique répond à l'objectif de regrouper les individus d'une population en classes formant une partition. C'est-à-dire que tout individu appartient à une classe et à une seule. La partition est l'ensemble des classes.

Différentes approches de classification

Il existe plusieurs méthodes de classification :

- effectuer le tri à plat ou croisé d'une ou plusieurs variables fermées uniques. Chaque classe est formée des individus ayant en commun une modalité (tri à plat) ou un couple de modalités (tri croisé). Toute variable fermée unique (ou toute combinaison de variables fermées uniques) définit ainsi une classification ;

- construire une typologie à vue à partir de l'examen d'un plan factoriel : on définit la variable fermée sur la typologie qui, comme toute variable fermée, définit une classification.

Dans le premier cas, la classification est la traduction exacte des données. Mais si on veut la construire à partir de plusieurs variables, on risque d'obtenir un nombre beaucoup trop grand de classes, ce qui fait perdre tout intérêt à la classification. Dans le deuxième cas, on peut fixer un nombre réduit de classes ou types, pour capter les individus apparaissant dans un plan factoriel mais le procédé d'affectation peut paraître trop subjectif et approximatif.

La procédure de classification automatique proposée est un compromis pour éviter les inconvénients qui viennent d'être signalés. Elle repose sur la mise en œuvre de la méthode dite des centres mobiles ou des nuées dynamiques.

Il existe bien d'autres méthodes (classifications hiérarchiques ascendantes ou descendantes) mais celle que nous proposons peut s'appliquer à de très grands effectifs et permet de contrôler a priori le nombre de classes.

Définir une classification

Les étapes de la méthode sont les suivantes :

- Fixer les objectifs de la classification ;

- Définir la population : tous les individus de la base ou une strate seulement ;

- Sélectionner les variables par rapport auxquelles on souhaite définir une partition de la population ;

- Fixer le nombre de classes désirées ou une partition de départ définie par la variable fermée unique de son choix ;

Il est tout à fait possible de choisir pour variable de départ une typologie définie à partir d'une analyse en composantes principales ou une analyse factorielle multiple. On pourra ainsi affiner les agrégations effectuées dans le plan factoriel.

La méthode des centres mobiles

Cet algorithme met en œuvre le principe suivant : on améliore progressivement la partition de départ en calculant, pour chaque classe, un *individu fictif moyen*. On définit alors une nouvelle classification, par agrégation aux individus fictifs moyens résultant de la classification précédente. On répète l'opération tant que la nouvelle classification diffère de la précédente.

L'individu fictif moyen est calculé comme le barycentre de la classe, l'agrégation se fait en fonction de la moindre distance.

Analyser la classification obtenue

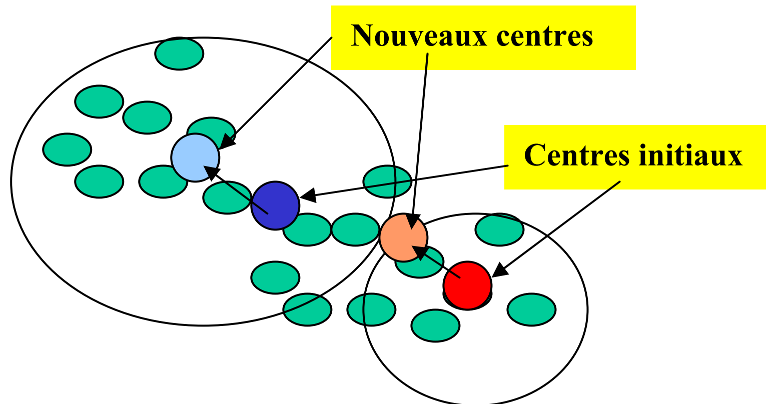
La classification résultant de l'algorithme dépend de la partition de départ. Elle est fixée d'une manière aléatoire si on se borne à indiquer le nombre de classes désirées, sinon, elle dépend de la variable choisie. Il peut alors être intéressant de rechercher plusieurs classifications pour les comparer entre elles.

On peut pour cela considérer :

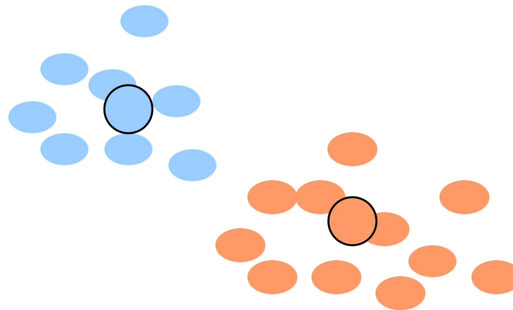
- la répartition des effectifs conduisant à des classes plus ou moins équilibrées ;

- le pouvoir discriminant de la classification. On cherchera pour cela à comparer les classes du point de vue des variables à partir desquelles on les a déterminées : analyse de la variance ou test du Chi² selon qu'il s'agit de variables quantitatives ou qualitatives.

La méthode des centres mobiles



Classification finale



Sommaire détaillé